ECOPHYSIOLOGY AND PHYLOGENY OF FAECALIBACTERIUM PRAUSNITZII IN HEALTHY AND DISEASED GUT. APPLICATION IN INFLAMMATORY BOWEL DISEASE DIAGNOSTICS

Mireia López Siles

Universitat de Girona

doctoral thesis

Ecophysiology and phylogeny of *Faecalibacterium prausnitzii* in healthy and diseased gut. Application in Inflammatory Bowel Disease diagnostics

mireia lópez siles
2015

Universitat de Girona

doctoral thesis

Ecophysiology and phylogeny of *Faecalibacterium prausnitzii* in healthy and diseased gut. Application in Inflammatory Bowel Disease diagnostics

mireia lópez siles
2015

Doctorate program in Experimental Sciences and Sustainability.

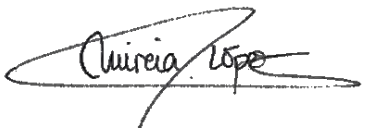
Thesis supervisors

Prof. Dr. L. Jesús García Gil
Professor in Microbiology

Dr. Margarita Martínez Medina
Lecturer in Microbiology

PhD candidate

Mireia López Siles



This thesis is submitted in fulfilment of the requirements to obtain the doctoral degree from the Universitat de Girona

Universitat de Girona

Hereby, Prof. Dr. L. Jesús García Gil and Dr. Margarita Martínez-Medina, of the Universitat de Girona,

CERTIFY:

That this doctoral thesis entitled **"Ecophysiology and phylogeny of *Faecalibacterium prausnitzii* in healthy and diseased gut. Application in Inflammatory Bowel Disease diagnostics."**, that Mireia López Siles has submitted to obtain the doctoral degree from the Universitat de Girona has been completed under their supervision, and meets the requirements to opt for the *International Doctor* mention.

In witness whereof and for such purposes as may arise, the following certification is signed:

Prof. Dr. L. Jesús García Gil

Professor in Microbiology

Dr. Margarita Martínez Medina

Lecturer in Microbiology

Girona, 2015

"En la vida siempre hay cosas demasiado complicadas para explicarlas en cualquier idioma" (…) "Difíciles de explicar no sólo a los demás, sino también a uno mismo. (…) En cualquier caso, mañana se aclararán muchas cosas. Sólo tienes que esperar. Y si no se aclara nada, qué se le va a hacer. Habrás hecho todo lo que estaba en tus manos."

Los años de peregrinación del chico sin color, Haruki Murakami

Dedicatòria

Això quedarà una mica llarg, però ja no vindrà d'unes quantes pàgines! Sens dubte fer la tesi no ha estat només una part més durant la meva formació acadèmica, sinó una part molt important de la meva formació com a persona. Molts sou els durant aquests anys heu fet possible tenir avui aquest llibret entre les mans, i el mínim que puc fer és dedicar-vos-hi unes paraules. Possiblement siguin les línies més difícils d'escriure de tota la Tesi. La pressió per no deixar-se ningú és gran. De tota manera, si ets tu...em sap greu però la memòria a vegades falla! Més enllà d'unes línies en aquests fulls, espero haver sabut agrair l'ajuda rebuda en el seu moment a tots els que m'heu acompanyat en aquest projecte.

En primer lloc, donar les gràcies a en **Jesús** per l'oportunitat d'entrar al món de la recerca, que tant m'apassiona. Gràcies per permetre'm aprendre tant durant aquest anys. **Marga**, merci per ser un punt de referència i de suport constant. Per creure en que acabaria més que jo mateixa, i ajudar-me a tirar endavant tant a nivell científic com personal, si he après tant durant aquests anys ha estat gràcies a tu.

També voldria donar les gràcies als companys del grup de Microbiologia de la Malaltia Intestinal amb qui he compartit experiències durant aquests anys: a la **Laia C**, per ajudar-me a donar les primeres passes en qPCR, part fonamental d'aquesta Tesi!; a la **Núria P** per la introducció al laboratori; a la **Tere** per iniciar els estudis amb pacients de càncer de colon que han fet aquest treball més complert; a en **Romà**, per la dedicació i l'entusiasme en els estudis dels filogrups;a en **Carles**, per tots els bons consells fruit de la seva experiència; i a la **Carla** i la **Txell** per tot el suport i energia positiva que m'han donat en la "recta final".

Igual d'importants han estat durant aquests anys la resta de companys de l'Àrea de "*Micro*": **Bo**, gràcies per totes les recomanacions literàries tant científiques com no, han fet que aquesta Tesi sigui una mica més rica; **Catxo**, gràcies per motivar-me a tenir sempre alguna cosa "al limbo"; **Frederic**, merci pel teu optimisme i per ser el meu mentor en les primeres passes en el món de la docència; **Xavi V**, gràcies per endinsar-me en el món de la Bioremediació; **Xevi T** moltes gràcies per ser un exemple constant de paciència i de les coses

ben fetes; **Geno**, moltes gràcies per preocupar-te sempre pel meu benestar; **Laia M** moltíssimes gràcies per estar sempre al peu del canó, disposada a donar un cop de mà quan faci falta i pel que faci falta: comandes, material, i un infinit etc. Sense dubte, els peque-Micros, han estat també MACRO-fonamentals: **Anna**, t'hauria de donar les gràcies per tantes coses i tantes vegades, que no acabaria mai! Ho deixarem en un gràcies per TOT!!!; **Rosalia** merci per fer-me de germaneta gran a nivell científic i personal!; **Arantxa** moltes gràcies pel teu humor, saps com arrencar sempre un somriure als que t'envolten, fins i tot estant al sud dels països catalans; **Lairos**, merci per tots els cops que m'has salvat, ja sigui d'una caiguda o d'un comentari d'un *referee* poc amable; **Àlex** gràcies per culturitzar-me amb les múltiples recomanacions musicals i de cine; **Ari** merci per estar disposada a donar un cop de mà ja sigui divulgant ciència o fent maratons per arreglar "el ganxo"!; **Sara** gràcies per les cançons que amenitzen les estones d'ordinador; **Elena** merci per ser una gran veïna, font inesgotable de tè, queviures i converses regeneradores. **Gela** moltes gràcies pel cop de mà amb la preparació de les pràctiques, una experiència que sens dubte ens va curtir!; **Imma** gràcies per totes les paraules amables i per estar disposada a fer costat "sense preguntes"; i també gràcies a l' **Olga**, la **Mireia F**, la **Jessica** i en **Sergi**, que malgrat estar a l'altra banda del turonet sempre és un plaer compartir dinars i tions amb vosaltres.

Han estat també molt importants en aquesta etapa la resta de personal investigador en formació (i molts ja doctors) amb qui he pogut compartir moments memorables: **Jess**, moltes gràcies per ser capaç de fer dolços els moments més amargs; **Montse**, gràcies per tenir el do de despertar-me el bon humor fins i tot els dies que no se si l'he deixat a casa!; **Pere**, merci per estar sempre a un "clic" i disposat a ajudar-me; **Santi**, gracias por estar siempre "al pie de la escalera-meeting point"; **Sílvia B** merci per les estones de desconnexió a la piscina i ajudar-me amb el maqueteig dels pdf; **Txell**, un plaer compartir congressos amb tu; **Dolors, Eva, Clara** i **Marta**, merci per les múltiples trobadetes aquests darrers mesos; **Olga**, merci per tots els bons moments compartits a EECC; **Marçal**, gràcies per adoptar el "sakazakii project", un plaer fer recerca amb tu; **Roger V** merci pour tous les notes, je n'aurai pas réussi ainsi à Français sans toi; **Pau**, gràcies per totes les coses que he après al teu costat, des de fer *heatmaps* fins que hi ha una fleca nova al barri!; **Sandra** merci per fer dels 30 una festa inoblidable; **Alexandra**, merci per ser professora meva durant i després de la carrera; m'encanta el teu riure contagiós!; **David D** me han encantado nuestras conversaciones infinitas a las 8 de la tarde; **Luís**, sempre saps com fer-me riure un munt! Viva!!!; i gràcies també a la **Cristina**, en **David**, la **Glòria**, la **Vero**, la **Mariona**, l'**Anna V**, la **Marta**, la **Vicky**, la **Laia,** en **Roger**, per tots els moments-becarils compartits, són la xispa del dia a dia! Ah, i a en **Pau C**, en **Dani** i en **Pedro**, perquè és bonic veure cada dia l'entusiasme i la il·lusió amb la que es comença al món de la recerca.

Gràcies també a la **Mercè F,** per sempre tenir la cadira confident a punt; i a tots els companys del Departament amb qui he pogut compartir docència o que han estat oberts a ajudar-me sempre que ho necessitat. Fent memòria t'adones que ha valgut molt la pena fer la tesi, i que he conegut gent molt maca!!!

Gràcies als doctors **Xavier Aldeguer** i **David Busquets** del Hospital Universitari Dr. Josep Trueta, i als doctors **Carles López Oliu** i **Míriam Sàbat** del IAS Santa Caterina per la seva col·laboració en la recollida de mostres (imprescindible per a tirar endavant aquest projecte) i per la seva ajuda i paciència per apropar-me al món mèdic i resoldre els múltiples dubtes que m'han sorgit durant el camí.

I would like to thank also those who hosted me abroad. **Harry**, many thanks for opening the doors of your lab. **Sylvia**, many thanks for taking care of me during my stay in Aberdeen but also afterwards, it has been a pleasure to work with and learn from you. I am grateful also to **Petra**, **Karen**, **Freda**, **Jenny**, **Rai**, **Ania**, **Masa**, **Dinesh, Álvaro** and **Eddy**, for making me feel more than wellcome at the Rowett. Y muchísimas gracias a **Eva**, **Cristina**, **Luz**, **Bea**, **David** y **Antonio**, por acogerme como una más del grupo y descubrirme las maravillas de Escocia, que son muchas!!! I appreciate also Dr. **M.Tanweer Khan** and **Dr. H.J.M. Harmsen** collaboration for this Thesis accomplishment.

Gràcies a la **Marta M**, perquè a part d'una molt bona companya de pis, s'ha convertit en una molt bona amiga; i també gràcies a la **Regina**, la **Marta H**, i la **Jes**, per tots els moments de convivència compartits com "marqueses", que no van ser pocs durant deu anys!!! Gràcies a la **Ruth**, la **Laura**, la **Silvia**, la **Txell**, la **Cristina** i els respectius, per totes les activitats que organitzem de tant en tant i que m'han ajudat a trobar moments de desconnexió quan més ho he necessitat. Merci també a la **Nuri P** i la **Laia B** per les nostres re-trobades. I sobretot gràcies a l'**Anna** i la **Nuri,** perquè amigues com vosaltres només es troben un cop a la vida, i jo he estat de sort perquè n'he trobat dues!!!

Por último, muchas gracias a mi familia y especialmente a mis padres. Gracias por toda la comprensión que habeis mostrado durante estos años. Gracias por vuestro cariño incondicional. Esta tesis es sobre todo gracias a vosotros.

Acknowledgments

Sources of projects funding

This work was partially founded by the Spanish Ministry of Education and Science through the projects:

- "BACTECCU: *Estudio de la microbiota asociada a la Enfermedad Inflamatoria Intestinal para la obtencion de marcadores bacterianos específicos: aplicación al diagnóstico de la Enfermedad de Crohn y Colitis Ulcerosa* **(SAF2006-00414)**, whose PI was Prof. L. Jesús García-Gil

- "GENAIEC: *Adherent Invasive Escherichia coli (AIEC): distribution among intestinal diseases other than Crohn's disease and genes involved in its pathogenicity*" **(SAF2010-15896)** whose PI was Prof. L. Jesús García-Gil

- "IDAIEC: *Determinantes genéticos implicados en el patotipo Escherichia coli adherente invasiva (AIEC) para su identificación en la enfermedad de Crohn*" **(SAF2013-43284P)** whose PI was Dr. Margarita Martínez Medina

PhD candidate fellowships

During the PhD period, Mireia López Siles was awarded with:

- FI pre-doctoral grant from the Generalitat de Catalunya **(2008FI_B 00008, 2009FI_B1 00001, 2010FI_B2 00135)**, that receives support from the European Union Comissionate

- A mobility fellowship from the University of Girona (call "*Ajuts complementaris per a la mobilitat d'investigadors de la Universitat de Girona 2009*") to support a research stay at the Rowett Institute of Nutrition and Health (from 14/09/2009 to 17/12/2009)

- A fellowship from the Ministry of Education and Science within the programme "*Subvenciones para estancias de movilitat de estudiantes para la obtención de la Mención Europea en el Título de Doctor*" (call 2009-2010) to support a research stays at the Rowett Institute of Nutrition and Health (from 22/06/2010 to 30/09/2010).

List of publications

The results of this PhD Thesis have been published in scientific journals included in the Journal Citation Report of the Institute of Scientific Information.

- Mireia Lopez-Siles, Tanweer M. Khan, Sylvia H. Duncan, Hermie J. M. Harmsen, L. Jesús Garcia-Gil, and Harry J. Flint.
  Cultured representatives of two major phylogroups of human colonic *Faecalibacterium prausnitzii* can utilize pectin, uronic acids and host-derived substrates for growth.
  Applied and Environmental Microbiology 2012; 78 (2): 420-428.
  IF_{2012}=3.778; Quartile 1 in categories Biotechnology&Applied Microbiology (32/160) and Microbiology (24/107).

- Mireia Lopez-Siles, Margarita Martinez-Medina, Carles Abellà, David Busquets, Miriam Sabat-Mir, Sylvia H. Duncan, Xavier Aldeguer, Harry J. Flint, and L. Jesús Garcia-Gil.
  Mucosa-associated *Faecalibacterium prausnitzii* population richness is reduced in inflammatory bowel disease patients.
  Applied and Environmental Microbiology 2015; 81 (21): 7582-7592.
  IF_{2014}= 3.668; Quartile 1 in categories Biotechnology&Applied Microbiology (34/162) and Microbiology (29/119).

- Mireia Lopez-Siles, Margarita Martinez-Medina, David Busquets, Miriam Sabat-Mir, Sylvia H. Duncan, Harry J. Flint, Xavier Aldeguer, and L. Jesús Garcia-Gil.
  Mucosa-associated *Faecalibacterium prausnitzii* and *Escherichia coli* co-abundance can distinguish Irritable Bowel Syndrome and Inflammatory Bowel Disease phenotypes.
  International Journal of Medical Microbiology 2014; 304:464-475.
  IF_{2014}=3.614; Quartile 2 in categories Microbiology (30/119) and Virology (11/33).

- Mireia Lopez-Siles, Margarita Martinez-Medina, Romà Surís-Valls, Xavier Aldeguer, Miriam Sabat-Mir, Sylvia H. Duncan, Harry J. Flint, and L. Jesús Garcia-Gil.
  Changes in the abundance of *Faecalibacterium prausnitzii* phylogroups I and II in the intestinal mucosa of inflammatory bowel disease and patients with colorectal cancer. *Accepted in Inflammatory Bowel Diseases on 3rd August 2015, doi:10.1097/MIB.0000000000000590*
  IF_{2014}=4.464; Quartile 1 in category Gastroenterology & Hepathology (15/76).

Patent

Part of the results of this PhD Thesis are included in a European patent application filed at the Spanish Patent and Trademarks Office (OEPM).

- Mireia Lopez-Siles, L. Jesús Garcia-Gil, Xavier Aldeguer, Margarita Martinez-Medina. **Method for the detection, follow up and/or classification of intestinal diseases.** Universitat de Girona, Institut d' Investigació Biomèdica de Girona Dr. Josep Trueta, Goodgut S.L. (Application number: EP15382427, Filling date: 11 August 2015)

This is a patent application that it is currently under examination.

Table of abbreviations

Abbreviation	Description
16S rRNA	Small subunit ribosomal gene
A1	CD diagnosed < 16 years old
A2	CD diagnosed at 17-40 years old
A3	CD diagnosed > 41 years old
AIEC	Adherent Invasive *Escherichia coli*
AN	Nutrient Agar
ANI	Average nucleotide identity
ASCAs	Anti-*Saccharomyces cerevisiae* antibodies
ATCC	American type culture collection
AUC	Area under the ROC curve
B1	Non-stricturing, non-penetrating CD
B2	Stricturing CD
B3	Penetrating CD
BA	Blood Agar
BHI	Brain Heart Infusion broth medium
bp	Base pair
C-CD	Colonic Crohn's disease
CD	Crohn's disease
CDAI	Crohn's disease activity index
CECT	Colección Española de Cultivos Tipo
cnPCR	Conventional polymerase chain reaction
Cq	Quantification cycle
CRC	Colorectal cancer
CRP	C-reactive protein
CZ	Colby and Zathman medium
DDH	DNA-DNA hybridization
dDDH	Digital DNA-DNA hybridization
DGGE	denaturing gradient gel electrophoresis
DNA	Deoxyribonucleic acid
dNTPs	Deosyribonucleoside triphosphates
DSMZ	Deutsche Sammlung von Mikroorganismen und Zellkulturen (German collection of microorganisms)
DSS	Dextran Sodium Sulfate
E1	Ulcerative proctitis
E2	Distal or left-sided ulcerative colitis
E3	Pancolitis or universal colitis
EDTA	Ethylenediaminetetraacetic acid
EMBL	European Molecular Biology Laboratory
F-E	*Faecalibacterium prausnitzii – Escherichia coli*
FAM™	6-carboxyfluorescin
FAP	Familial adenomatous polyposis
FISH	Fluorescent *in situ* hybridisation
FP	*Faecalibacterium prausnitzii*
G+C	Guanine and Cytosine content
H	Healthy subjects
H'	Shannon diversity index
IAC	Internal Amplification Control
IBD	Inflammatory bowel disease
IBS	Irritable bowel syndrome
IBS-C	Constipation-predominant irritable bowel syndrome
IBS-D	Diarrhoea-predominant irritable bowel syndrome
IBS-M	Alternating irritable bowel syndrome
IC	Ischemic colitis
IC-CD	Ileocolonic Crohn's disease
I-CD	Ileal Crohn's disease
IdC	Indeterminate colitis
IFN-γ	Interferon gamma
IL-8	Interleukin-10 or chemokine (C-X-C motif) ligand 8 (CXCL8)
IL-10	Interleukin-10 (IL-10), also known as human cytokine synthesis inhibitory factor (CSIF)
IL-12	Interleukin-12
JC	Jukes-Cantor algorithm for phylogenetic distance analysis
kDa	Kilodaltons
L1	Ileal-CD
L2	Colonic CD

Abbreviation	Description
L3	Ileocolonic CD
L4	CD affecting upper gastrointestinal tract
LiB	Liver Broth
M2GSC	Modified Med2 of Hobson (1969) and modified by Miyazaki *et al*. (1997)
MIC	Minimum inhibitory concentration
µg	Microgram
µl	Microliter
µM	Micromolar
mg	Milligram
ml	Millilitre
mmHg	Millimeter of mercury
mM	Millimolar
MRS	Man, Rogosa and Sharpe medium
MUM	Maximal unique matches
mV	Millivolts
N	Sample size
NCBI	National Center for Biotechnology Information
NCTC	National Collection of Type Cultures
ND	Not determined
NET	Neuroendocrine tumour of the midgut
NF-κB	NF-κB (nuclear factor kappa-light-chain-enhancer of activated B cells)
ng	Nanogram
NGS	Next Generation Sequencing
NJ	Neighbour-Joining
nm	Nanometre
NOD2	Nucleotide-binding oligomerization domain-containing protein 2 (NOD2) also known as caspase recruitment domain-containing protein 15 (CARD15) or inflammatory bowel disease protein 1 (IBD1)
OD	Optical Density (usually accompanied by a subscript indicating wavelength in nm)
OTU	Operational taxonomic unit
pANCAs	Perinuclear anti-neutrophilic cytoplasmatic antibodies
PCR	polymerase chain reaction
PCR-DGGE	Polymerase chain reaction-denaturing gradient gel electrophoresis
PDA	Potato Dextrose agar
PHG I	*Faecalibacterium prausnitzii* phylogroup I
PHG I-E	*Faecalibacterium prausnitzii* phylogroup I-*Escherichia coli*
PHG II	*Faecalibacterium prausnitzii* phylogroup II
PHG II-E	*Faecalibacterium prausnitzii* phylogroup II-*Escherichia coli*
PPARγ	Peroxisome proliferator-activated receptor gamma
qPCR	Quantitative real-time polymerase chain reaction
RAPD-PCR	Random Amplified Polymorphic DNA PCR
RDP	Ribosomal Database Project
RNA	Ribonucleic acid
RNAse	Ribonuclease
ROC	Receiver Operating Characteristic
rRNA	Ribosomal ribonucleic acid
SCFA	Short Chain Fatty Acids
SD	Standard deviation
SES-CD	Simple Endoscopic Score for CD
SLC	Self-limiting colitis
sp.	Species (singular)
spp.	Species (plural)
T1-T4	Tumour state for CRC patients
TAE	Tris-acetate-EDTA buffer
TAMRA™	Tetramethyl-6-Carboxyrhodamine
T-RFLP	Terminal restriction fragment length polymorphism
Tris-HCl	trisaminomethane hydrochloride
TNF-α	Tumour necrosis factor alpha
UC	Ulcerative colitis
UGC	Upper gastrointestinal cancer
UPGMA	Unweighted-pair group method using average linkages (used to built dendograms)
V	volts
VIC®	6-carboxyrhodamine
wt/vol	Weight/volume
YCFA	Yeast Casitone Fatty-Acids medium
YcFA	YCFA with casitone content reduced to 0.2%
YCFG	YCFA medium supplemented with 10mM glucose
yr	Year

List of figures

List of tables

Table of contents

Summary

The microbial community inhabiting human intestine plays a fundamental role for health. A rising number of studies have reported that patients suffering intestinal disorders feature an altered gut microbiota in comparison to healthy subjects. To gain knowledge about the diversity and functions of key members of the intestinal microbiota is essential to better understand the role of these symbionts to maintain human health. In addition, bacteria that correlate with healthy gut status can be identified among the bacterial community inhabiting our gut, thus offering a myriad of novel biomarkers to assess intestinal health and monitor gut disorders. Finally interventions aimed at modulating gut microbiota in order to restore a healthy-like community, offers a novel strategy to achieve human health.

Faecalibacterium prausnitzii, a member of the phylum Firmicutes (Ruminococcaceae), is one of the three most abundant species found in the human gut. Interest in this species has increased in the last years since it was reported that *F. prausnitzii* is depleted in patients suffering inflammatory bowel diseases (IBD), thus pointing out its beneficial role to maintain gut health. However, little information about its growth requirements, the genetic diversity comprised within this species, and how its abundance is affected by intestinal disorders was reported. Therefore the main purpose of this work was to gain insight into this species physiology, diversity and abundance in healthy and diseased gut.

To achieve this objective, first a phylogenetic and phenotypic characterisation of *F. prausnitzii* isolates was performed in order to determine which carbon sources found in the gut can be used for *F. prausnitzii* to grow, and to assess its sensitivity to changes in gut environmental factors (**Chapter 1**). Phylogenetic analysis based on the 16S rRNA sequences indicated that the available isolates separate into two phylogroups which have a 97% of this gene similarity. *F. prausnitzii* isolates were metabolically versatile, capable to grow on carbohydrates of different structure and origin (host- and diet-derived substrates). All strains tested were bile-sensitive, showing at least 80% growth inhibition in the presence of 0.5 % (wt/vol) bile salts, while inhibition at mildly acidic pH was strain dependent. These attributes help to explain the abundance of *F. prausnitzii* in the colonic community, but also suggest factors in the gut environment that may limit its presence in a diseased gut.

Since gut environmental conditions are different between a healthy and a diseased intestine, the second part of this work was aimed at determining if subjects with gastrointestinal disease host in the colon different mucosa-associated *F. prausnitzii* populations from healthy in terms of richness and composition (**Chapter 2**). A novel species-specific polymerase chain reaction-denaturing gradient gel electrophoresis (PCR-

DGGE) targeting the 16S rRNA gene was developed to fingerprint *F. prausnitzii* populations in biopsies from healthy subjects (H) and patients suffering intestinal disorders such as irritable bowel syndrome (IBS), ulcerative colitis (UC), Crohn's disease (CD) and colorectal cancer (CRC). The richness of *F. prausnitzii* subtypes was lower in IBD patients than in H subjects. The most prevalent operational taxonomic units (OTU) were shared by all the patients groups, but their distribution and the presence of some disease-specific *F. prausnitzii* phylotypes allowed differentiating IBD and CRC population from that in H. This prompted further studies to address the suitability of their quantification as putative biomarkers of disease.

Therefore, in the third part of this work it was explored the usefulness of *F. prausnitzii* quantification to assist in either gut disorders diagnostic or prognostic (**Chapter 3**). The load of this species was determined by novel quantitative polymerase chain reaction (qPCR) assays in ileal, colonic and rectal biopsy samples of H, IBS, IBD and CRC subjects.

On the one hand it was explored the usefulness of total *F. prausnitzii* as biomarker in conjunction with *Escherichia coli* (an other extensively reported representative of IBD dysbiosis), and the *F. prausnitzii-E. coli* index (F-E index) was calculated. IBD patients had lower *F. prausnitzii* abundance than H and IBS. CD patients showed higher *E. coli* counts than H and UC patients. The F-E index discriminated between H, CD and UC patients, and even between disease phenotypes that are usually difficult to distinguish as ileal-CD (I-CD) from ileocolonic-CD and colonic-CD (C-CD) from extensive colitis (E3). *E. coli* increased in active CD patients, and remission in I-CD patients was compromised by high abundance of this species. Treatment with anti-tumor necrosis factor (TNF) α diminished *E. coli* abundance in I-CD whereas none of the treatments counterbalanced *F. prausnitzii* depletion. These results demonstrate that *F. prausnitzii* and *E. coli* are useful indicators to assist in IBD phenotype classification. In addition, the abundance of these species could also be used as a supporting prognostic tool in I-CD patients. Our data indicates that current medication does not restore these two species levels to those found in a healthy gut, and yet more precise biomarkers should be found to discriminate between some subtypes of IBD.

On the other hand, it was assess the suitability of *F. prausnitzii* phylogroups depletion as biomarkers for gut diseases. Lower levels of phylogroup I were found in CD, UC and CRC compared with H subjects. Phylogroup I load was a better biomarker than total *F. prausnitzii* to discriminate subjects with gut disorders from H. Phylogroup II depletion was observed only in CD patients, and can be potentially applied to differentiate E3 from C-CD. Phylogroup I was lower in active CD patients whereas those CD with intestinal resection showed a reduction in phylogroup II. Treatments with mesalazine and immunosupressants did not result in the recovery of *F. prausnitzii* phylogroups abundance. Quantification of

F. prausnitzii phylogroups may help to precisely identify gut disorders, and to classify IBD location.

The results of this work provide new insights to understand *F. prausnitzii* physiology and its distribution in the gut. Novel evidences of this species population alterations found in diseased gut have been revealed. Furthermore, the results are in agreement with previous research on the microbial community of patients suffering intestinal disorders, which indicate that this species is depleted in gut abnormal conditions. The present study gives some clues to reveal possible causes. Finally, novel molecular tools have been provided, and its usefulness to discriminate between conditions has been proven, thus implying a step forward in the field of intestinal disorders diagnostics.

Resum

La comunitat microbiana que habita el tracte intestinal humà juga un rol fonamental per a la salut. Nombrosos estudis han evidenciat que les persones que pateixen malalties intestinals presenten una microbiota intestinal alterada en comparació amb les persones sanes. Conèixer la diversitat i funcions dels membres principals de la microbiota intestinal és essencial per comprendre millor el paper que juguen aquests simbionts per mantenir la salut de l'ésser humà. A més, es pot identificar quins dels microorganismes que constitueixen la comunitat microbiana de l'intestí correlacionen amb un estat sa de salut intestinal, oferint així la possibilitat d'identificar nous biomarcadors per avaluar l'estat de salut intestinal i monitorar l'evolució de malalties del tracte intestinal. Per últim, l'aplicació de intervencions nutricionals destinades a modular la microbiota intestinal a fi de restaurar una comunitat similar a la que es troba en individus sans ofereix una nova estratègia per millorar la salut humana.

Faecalibacterium prausnitzii*, un membre del fílum Firmicutes (Ruminococcaceae), és una de les tres espècies més abundants del tracte intestinal humà. L' interès en aquesta espècie ha augmentat en els darrers anys des que es va evidenciar que* F. prausnitzii *desapareix en pacients que pateixen malaltia inflamatòria intestinal (IBD), posant de manifest el seu rol beneficial per mantenir la salut intestinal. No obstant, existeix poca informació sobre quins requeriments nutricionals té aquest microorganisme, la diversitat genètica que s'inclou dins aquesta espècie i com la seva abundància es veu alterada en pacients que pateixen malalties de l'intestí. Aquest treball té com a objectiu principal comprendre millor la fisiologia, diversitat i abundància de* F. prausnitzii *en individus sans i pacients amb malaltia intestinal.*

*Per assolir aquest objectiu, en primer lloc es va realitzar una caracterització filogenètica i fenotípica dels aïllats, a fi de determinar quines fonts de carboni que es poden trobar a l'intestí són les que principalment utilitza aquesta espècie per créixer, i definir la seva sensibilitat a canvis en factors ambientals de l'intestí (**Capítol 1**). L'anàlisi filogenètica del gen del 16S rRNA va mostrar que les soques actualment aïllades de* F. prausnitzii *es divideixen en dos filogrups amb un 97% de similitud en la seqüència d'aquest gen. La caracterització fenotípica va revelar que* F. prausnitzii *és un bacteri metabòlicament versàtil, que pot créixer utilitzant substrats amb un grau de complexitat variable, ja siguin procedents de la dieta o de l'hoste. Totes les soques van ser extremadament sensibles a sals biliars, mostrant com a mínim un 80% d'inhibició del creixement en presència de 0,5 % (pes/volum) de sals biliars. En canvi, la sensibilitat a canvis en el pH del medi va resultar ésser variable en funció de cada soca. Aquestes característiques permeten explicar l'elevada abundància de* F. prausnitzii *en la comunitat microbiana del còlon. El fet que tots els representats cultivables de* F. prausnitzii *mostrin una elevada sensibilitat a petits canvis en les condicions ecològiques que s'espera que ocorrin en determinades malalties intestinals, seria una possible explicació pel fet que l'abundància d'aquest bacteri comensal es trobi compromesa en un còlon alterat.*

*Donat que les condicions ambientals de l'intestí varien entre un intestí sa i malalt, a la segona part d'aquest treball es va voler determinar si les persones que pateixen un trastorn gastrointestinal tenen una població de F. prausnitzii associada a la mucosa colònica diferent de la que presenten els individus sans a nivell de riquesa i composició (**Capítol 2**). Es va desenvolupar un nou sistema de reacció en cadena de la polimerasa-electroforesi amb gel amb gradient desnaturalitzant (PCR-DGGE) específic per aquesta espècie i dirigit al gen del 16S rRNA. Es va analitzar el perfil de la població de* F. prausnitzii *en biòpsies colòniques de persones sanes (H), i pacients amb trastorns intestinals com ara síndrome del budell irritable (IBS), colitis ulcerosa (UC), malaltia de Crohn (CD) i càncer colorectal (CRC). La riquesa de subtipus de* F. prausnitzii *va ser menor en pacients amb IBD que en individus H. Les unitats taxonòmiques operacionals (OTU) més prevalents es van detectar en tots els grups d'individus. No obstant, la seva distribució i la presència de filotips específics de cada malaltia va permetre diferenciar les poblacions de* F. prausnitzii *d'IBD i CRC respecte a les que es troben en H. Aquestes evidències van servir com a base per a la identificació de nous biomarcadors a quantificar amb l'objectiu d'assistir en la identificació d'estats de malaltia intestinal.*

Per tant, a la tercera part d'aquest treball es va explorar l'aplicació de quantificar F. prausnitzii *com a biomarcador d'ajuda al diagnòstic o pronòstic de malalties intestinals (**Capítol 3**). La quantitat de* F. prausnitzii *es va determinar mitjançant nous assajos de reacció en cadena de polimerasa quantitativa (qPCR) en biòpsies d'ili, colon i recte d'individus H, IBS, IBD i CRC.*

D'una banda, es va estudiar la utilitat de F. prausnitzii *com a biomarcador conjuntament amb la quantificació d'*Escherichia coli (*un altre microorganisme extensament descrit com a membre de la disbiosi que presenten els pacients amb IBD i es va calcular l'índex* F. prausnitzii-E. coli *(index F-E). Els pacients amb IBD van presentar una menor abundància de* F. prausnitzii *que els individus H i amb IBS. Els pacients amb CD van mostrar una major quantitat d'*E. coli *en comparació amb els individus H i UC. L'índex F-E va permetre discriminar entre H, CD i pacients amb CU. Aquest índex també va permetre diferenciar entre fenotips d'IBD que solen ser difícils de discriminar. Per exemple va permetre distingir entre CD d'afectació ileal (I-CD) i CD d'afectació ileo-colònica (IC-CD), i entre CD colònica (C-CD) i pacients amb colitis ulcerosa extensa (E3). Es va observar un augment d'* E. coli *en pacients amb CD activa, i que una elevada abundància d'aquesta espècie comprometia el temps de remissió en pacients amb I-CD. El tractament amb factor de necrosi anti-tumoral (TNF)* α *va permetre disminuir l'abundància d'*E. coli *en pacients amb I-CD mentre que cap dels tractaments va permetre contrarestar la disminució de* F. prausnitzii. *Aquests resultats demostren que* F. prausnitzii *i* E. coli *són bons indicadors per ajudar en la classificació de fenotips de IBD. A més, l'abundància d'aquestes espècies també podria ser utilitzada com a biomarcador de suport al pronòstic en pacients amb I-CD. Es va observar que la medicació actual no restaura els nivells d'aquestes dues espècies als valors que es troben en un intestí sa, i que cal cercar indicadors més precisos per discriminar entre alguns subtipus d'IBD.*

D'altra banda, es va avaluar la utilitat dels filogrups de F. prausnitzii *com a biomarcadors pel diagnòstic de malalties intestinals. Els pacients amb CD, UC i CRC presenten una menor quantitat del filogrup I en comparació amb els individus H. L'abundància del filogrup I va ser un millor biomarcador en comparació amb la quantitat total de* F. prausnitzii *per discriminar els individus H respecte els pacients amb trastorns intestinals. La disminució de filogrup II es va observar només en pacients amb CD i aquesta característica pot ésser aplicada per diferenciar pacients amb E3 d'aquells amb C-CD. L'abundància del filogrup I va disminuir en pacients amb CD activa, mentre que els pacients amb resecció intestinal van mostrar una reducció en la quantitat de filogrup II. Els tractaments amb mesalazina i immunosupressors no van permeten restaurar l'abundància de cap dels dos filogrups de* F. prausnitzii. *Aquestes dades evidencien que la quantificació dels filogrups de* F. prausnitzii *permet una millor discriminació entre trastorns intestinals, i subtipus de IBD.*

Aquest treball aporta noves dades que permeten entendre millor la fisiologia i distribució a l'intestí de F. prausnitzii. *A més, s'ha evidenciat per primer cop que les poblacions d'aquesta espècie estan alterades en situació de malaltia intestinal. Els resultats obtinguts concorden amb les dades prèvies sobre la comunitat microbiana de pacients que pateixen malalties intestinals, on ja s'havia indicat que aquesta espècie es troba disminuïda. El present treball permet dilucidar les possibles causes d'aquest fenomen. Finalment en aquest estudi s'han dissenyat i optimitzat noves eines moleculars, i s'ha comprovat la seva capacitat per discriminar entre trastorns intestinals, el que implica una estratègia prometedora per aplicar en un futur en el camp del diagnòstic de les malalties intestinals.*

Resumen

La comunidad microbiana que habita el tracto intestinal humano juega un rol fundamental para la salud. Numerosos estudios han evidenciado que las personas que padecen enfermedades intestinales tienen una microbiota intestinal alterada en comparación con las personas sanas. Conocer la diversidad y funciones de los miembros principales de la microbiota intestinal es esencial para comprender mejor el papel que juegan estos simbiontes para mantener la salud del ser humano. Además, se puede identificar cuáles de los microorganismos que constituyen la comunidad microbiana del intestino correlacionan con un estado sano de salud intestinal, ofreciendo así la posibilidad de identificar nuevos biomarcadores para evaluar el estado de salud intestinal y monitorizar la evolución de las enfermedades del tracto intestinal. Por último, la aplicación de intervenciones nutricionales destinadas a modular la microbiota intestinal, con el objetivo de restaurar una comunidad similar a la que se encuentra en individuos sanos, ofrece una nueva estrategia para mejorar la salud humana.

Faecalibacterium prausnitzii*, un miembro del filo Firmicutes (Ruminococcaceae), es una de las tres especies más abundantes del tracto intestinal humano. El interés en esta especie ha crecido en los últimos años desde que se evidenció que* F. prausnitzii *desaparece en pacientes que padecen enfermedad inflamatoria intestinal (IBD), poniendo de manifiesto su rol beneficial para mantener la salud intestinal. Sin embargo, existe poca información sobre qué requerimientos nutricionales tiene este microorganismo, la diversidad genética que se incluye dentro de esta especie y como su abundancia se ve alterada en pacientes que sufren enfermedades del intestino. Este trabajo tiene como objetivo principal comprender mejor la fisiología, diversidad y abundancia de* F. prausnitzii *en individuos sanos y pacientes con enfermedad intestinal.*

*Para alcanzar este objetivo, en primer lugar se realizó una caracterización filogenética y fenotípica de los aislados a fin de determinar cuáles son las principales fuentes de carbono que se pueden encontrar en el intestino utilizadas por esta especie para crecer, y definir su sensibilidad a cambios en factores ambientales del intestino (**Capítulo 1**). El análisis filogenético del gen del 16S rRNA mostró que las cepas actualmente aisladas de* F. prausnitzii *se dividen en dos filogrupos con un 97% de similitud en la secuencia de este gen. La caracterización fenotípica reveló que* F. prausnitzii *es una bacteria metabólicamente versátil, que puede crecer utilizando sustratos con un grado de complejidad variable, ya sean procedentes de la dieta o del huésped. Todas las cepas fueron extremadamente sensibles a sales biliares, mostrando al menos un 80% de inhibición del crecimiento en presencia de 0,5 % (peso / volumen) de sales biliares. En cambio, la sensibilidad a cambios en el pH del medio resultó ser variable en función de cada cepa. Estas características permiten explicar la elevada abundancia de* F. prausnitzii *en la comunidad microbiana del colon. El hecho de que todos los representados cultivables de* F. prausnitzii *muestren una elevada sensibilidad a pequeños cambios en las condiciones ecológicas que se espera que ocurran en determinadas enfermedades intestinales, sería una*

posible explicación para el hecho de que la abundancia de esta bacteria comensal se encuentre comprometida en un colon alterado.

Dado que las condiciones ambientales del intestino varían entre un intestino sano y enfermo, en la segunda parte de este trabajo se quiso determinar si las personas que sufren un trastorno gastrointestinal tienen una población de F. prausnitzii *asociada a la mucosa colónica diferente de la que presentan los individuos sanos a nivel de riqueza y composición (**Capítulo 2**). Se desarrolló un nuevo sistema de reacción en cadena de la polimerasa-electroforesis en gel con gradiente desnaturalizante (PCR-DGGE) específico para esta especie y dirigido al gen del 16S rRNA. Se analizó el perfil de la población de* F. prausnitzii *en biopsias colónicas de personas sanas (H), y pacientes con trastornos intestinales tales como síndrome del intestino irritable (IBS), colitis ulcerosa (UC), enfermedad de Crohn (CD) y cáncer colorrectal (CRC). La riqueza de subtipos de* F. prausnitzii *fue menor en pacientes con IBD que en individuos H. Las unidades taxonómicas operacionales (OTU) más prevalentes se detectaron en todos los grupos de individuos. No obstante, su distribución y la presencia de filotips específicos de cada enfermedad permitieron diferenciar las poblaciones de* F. prausnitzii *de IBD y CRC respecto a las halladas en H. Estas evidencias han sido la base para la identificación de nuevos biomarcadores a cuantificar con el objetivo de asistir en la identificación de estados de enfermedad intestinal.*

Por tanto, en la tercera parte de este trabajo se exploró la aplicación de cuantificar F. prausnitzii *como biomarcador de ayuda al diagnóstico o pronóstico de enfermedades intestinales (**Capítulo 3**). La cantidad de* F. prausnitzii *se determinó mediante nuevos ensayos de reacción en cadena de polimerasa cuantitativa (qPCR) en biopsias de íleon, colon y recto de individuos H, IBS, IBD y CRC.*

En primer lugar, se estudió la utilidad de F. prausnitzii *como biomarcador conjuntamente con la cuantificación de* Escherichia coli *(otro microorganismo extensamente descrito como miembro de la disbiosis que ocurre en IBD), y se calculó el índice* F. prausnitzii-E. coli *(índice F-E). Los pacientes con IBD presentaron una menor abundancia de* F. prausnitzii *que los individuos H y con IBS. Los pacientes con CD mostraron una mayor cantidad de* E. coli *en comparación con los individuos H y UC. El índice FE permitió discriminar entre H, CD y pacientes con CU. Este índice también permitió diferenciar entre fenotipos de IBD que suelen ser difíciles de discriminar. Por ejemplo permitió distinguir entre CD de afectación ileal (I-CD) y CD de afectación ileo-colónica (IC-CD), y entre CD colónica (C-CD) y pacientes con colitis ulcerosa extensa (E3). Se observó un aumento de* E. coli *en pacientes con CD activa, y que una elevada abundancia de esta especie comprometia el tiempo de remisión en pacientes con I-CD. El tratamiento con factor de necrosis anti-tumoral (TNF)* α *permitía disminuir la abundancia de* E. coli *en pacientes con I-CD, mientras que ninguno de los tratamientos permitió contrarrestar la disminución de* F. prausnitzii. *Estos resultados demostraron que* F. prausnitzii *y* E. coli *son buenos indicadores para ayudar en la clasificación de fenotipos de IBD. Además, la abundancia de estas especies también podría ser utilizada como biomarcador de apoyo al pronóstico en pacientes con I-CD. Se observó que la medicación actual no restaura*

los niveles de estas dos especies a los valores que se encuentran en un intestino sano, y que aún se requieren indicadores más precisos para discriminar entre algunos subtipos de IBD.

En segundo lugar, se evaluó la utilidad de los filogrupos de F. prausnitzii *como biomarcadores para el diagnóstico de enfermedades intestinales. Los pacientes con CD, UC y CRC presentan una menor cantidad de* F. prausnitzii *total y del filogrupo I en comparación con los individuos H. La abundancia del filogrupo I fue un mejor biomarcador en comparación con la cantidad total de* F. prausnitzii *para discriminar los individuos H respecto a los pacientes con trastornos intestinales. La disminución de filogrupo II se observó sólo en pacientes con CD y esta característica puede ser aplicada para diferenciar pacientes con E3 de aquellos con C-CD. La abundancia del filogrupo I disminuyó en pacientes con CD activa, mientras que los pacientes con resección intestinal mostraron una reducción en la cantidad de filogrupo II. Los tratamientos con mesalazina y inmunosupresores no permitieron restaurar la abundancia de ninguno de los dos filogrupos de* F. prausnitzii. *Estos datos evidencian que la cuantificación de los filogrupos de* F. prausnitzii *permite lograr una mejor discriminación entre trastornos intestinales, y subtipos de IBD.*

Este trabajo aporta nuevos datos que permiten entender mejor la fisiología y distribución en el intestino de F. prausnitzii. *Además, se ha evidenciado por primera vez que las poblaciones de esta especie están alteradas en situación de enfermedad intestinal. Los resultados obtenidos concuerdan con los datos previos sobre la comunidad microbiana de pacientes que padecen enfermedades intestinales, donde ya se había indicado que esta especie se encuentra disminuida. El presente trabajo permite dilucidar las posibles causas de este fenómeno. Finalmente en este estudio se han diseñado y optimizado nuevas herramientas moleculares, y se ha comprobado su capacidad para discriminar entre trastornos intestinales, lo que implica una estrategia prometedora para aplicar en un futuro en el campo del diagnóstico de las enfermedades intestinales.*

Preface

This PhD Thesis covers a series of phylogenetic and ecophysiological studies on the gut simbiont *Faecalibacterium prausnitzii*, and points out its usefulness as healthy-like microbiota biomarker.

In the **Scientific background**, the state of the art about the composition and metabolic activities of the gut microbiota is presented, paying careful attention to its establishment and changes through life, its composition, which factors modulate this community, and the mutualistic relationship between bacteria and the host. The main intestinal disorders studied in this Thesis are described with a focus on the imbalance that occurs in the gut microbial community. Finally *F. prausnitzii* phylogeny, physiology, interactions with the host, and changes in abundance in different intestinal disorders have been reviewed. This allowed setting the scenario and framing the **Scope and Aims** of this Thesis.

The **Results** of this Thesis are presented in form of four Articles following the format of the intended Journal of publication. Each Article comprises a brief Introduction, plus a Materials and Method, Results and Discussion sections. At the end of each Article specific references cited are listed. For coherence with the Aims proposed, these articles have been compiled into three Chapters. In the first chapter *(Article I)*, based on a classical microbiology approach, a phylogenetic and phenotypic characterisation of *F. prausnitzii* isolates was performed. This has allowed gaining insight into the ecophysiology of this species, and to evidence environmental factors of the human gut that may modulate its populations in diseased gut. In the second chapter *(Article II)*, by means of microbial ecology techniques, the colonic mucosa-associated *F. prausnitzii* populations in healthy and diseased gut have been profiled and differences in richness and composition have been revealed. In the third chapter (*Articles III* and *IV*), studies of applied microbiology are compiled. The abundance of mucosa-associated *F. prausnitzii* in several intestinal disorders has been assessed, and its usefulness as diagnostic biomarker in inflammatory bowel diseases has been explored.

A **Concluding remarks** section that integrates discussion about topics treated across all the Results chapters has been included. This section is intended to elucidate how this work has provided novel information about *F. prausnitzii* physiology, ecology and usefulness as biomarker. Besides, this section highlights the coherence behind the discrete pieces of work presented in the Results, and leads the path towards **Conclusions**.

Lastly, a section with **References** (except those specifically detailed in the Articles), an **Annex** with Supplemental Materials of the Articles and a **Glossary** listing technical terms that appear throughout the Thesis have been included.

1 scientific background

1.1. Composition and metabolic activity of gut microbiota

The composition, activity and significance for the host of the gut microbiota have been scantly known until recently because of its complexity, the difficulty of accurate sampling and limitations on the available techniques. However, the knowledge in the field of intestinal microbiota has evolved at an unprecedented rate in the last years, gaining insight into its essential role in human health and disease.

1.1.1. Colonisation of gut microbiota and changes through life stages

The establishment of gut microbiota starts immediately after birth when the gastrointestinal tract of a newborn is rapidly colonised by microorganisms from the mother and the surrounding environment. Mode of delivery and feeding regime play a crucial role in shaping the acquisition and the structure of the microbiota in neonates [1, 2]. Gut microbiota of vaginally-delivered babies resembles maternal microbiota, whereas it can differ significantly in those delivered via Caesarean section.

Bifidobacteria predominate in breast-fed babies while the population found in bottle-fed babies is more diverse, and dominated by taxa such as Bacteroidetes in detriment of Bifidobacteria (Figure 1.1) [2, 3]. During this initial phase the microbial composition of the gut shows low diversity, richness and evenness, and it is extremely variable [4]. It has been estimated that it is only after the introduction of solid food when the gut microbiota stabilizes

and progressively develops towards that found in young adults (section 1.1.2), with increased diversity and abundance of anaerobic Firmicutes [5]. A recent study using high-throughput microarray analysis showed that the establishment of an adult-like intestinal microbiota occurs at a later age than previously reported, as in four-year old children the microbiota has not yet reached the climax of bacterial diversity [4].

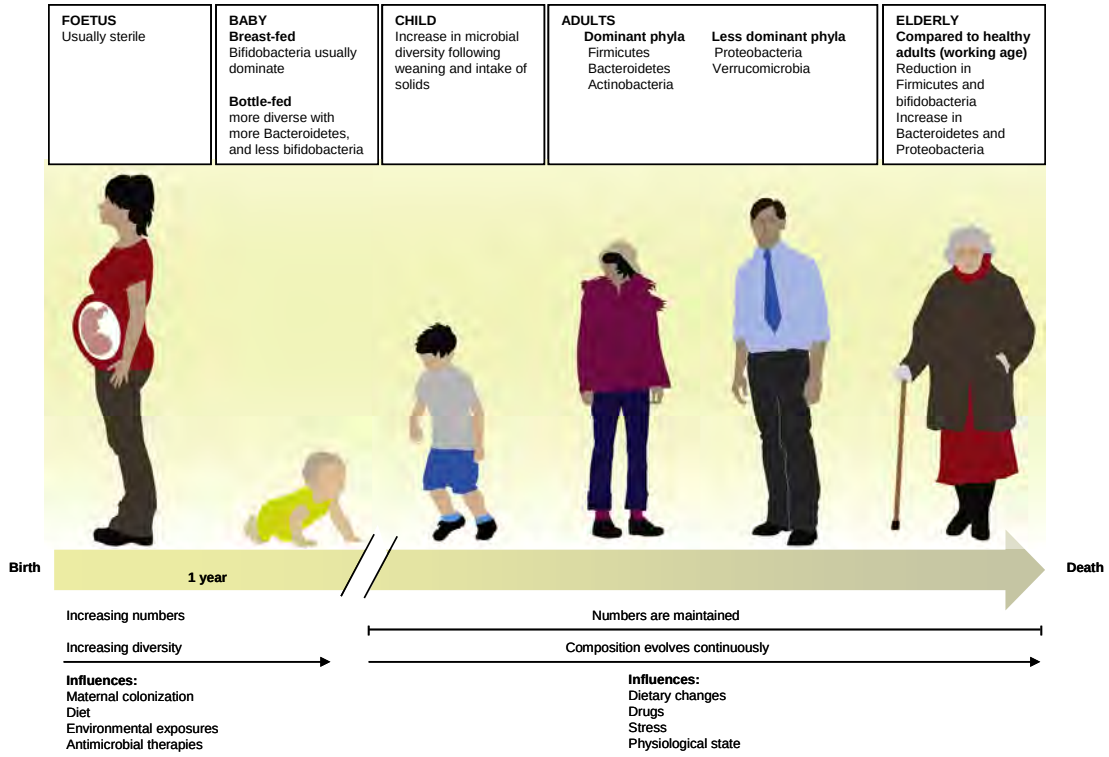


Figure 1.1 Summary of the changes in the gut microbiota composition through life time (adapted from [6, 7]).

The intestinal microbiota of an individual remains rather host-specific throughout life, although many factors such as diet, medication, stress and age amongst others can influence it (extended in section 1.1.3). It is noteworthy that gut microbiota has been reported to be less diverse in elderly, and mostly characterised by a reduction in Firmicutes and bifidobacteria, with a concomitant increase in Bacteroidetes and Enterobacteriaceae (for review see [6, 7]).

1.1.2. Composition of gut microbiota

The human gut is the preferred site for colonisation of microorganisms as it is characterised by a large surface and a constant input of molecules and substances (either incorporated through diet or produced by the host) which can be used as nutrients. It has been estimated that the gut microbiota of a healthy human adult is a complex ecological ecosystem consisting of approximately 10^{14} microorganisms. This population is the largest microbial community associated with the human body and outnumbers by a factor of 10 the number of human cells [7].

However, gut microbiota is not homogeneously distributed along the gastrointestinal tract. The average microbial abundance in the stomach, small intestine and colon has been estimated to be ~10^1, ~10^7 and ~10^{12} cells per gram of content respectively (Figure 1.2). However, these numbers have been obtained in studies based on culture methods, and although they are extensively used and accepted by the scientific community, different counts have been found using molecular approaches. Surprisingly, higher bacterial concentrations have been found in terminal ileum mucosa (1.3×10^{12} 16S rRNA gene copies per g of mucosal tissue) than in colonic mucosa (2.5×10^{10} 16S rRNA gene copies per g of mucosal tissue) [8].

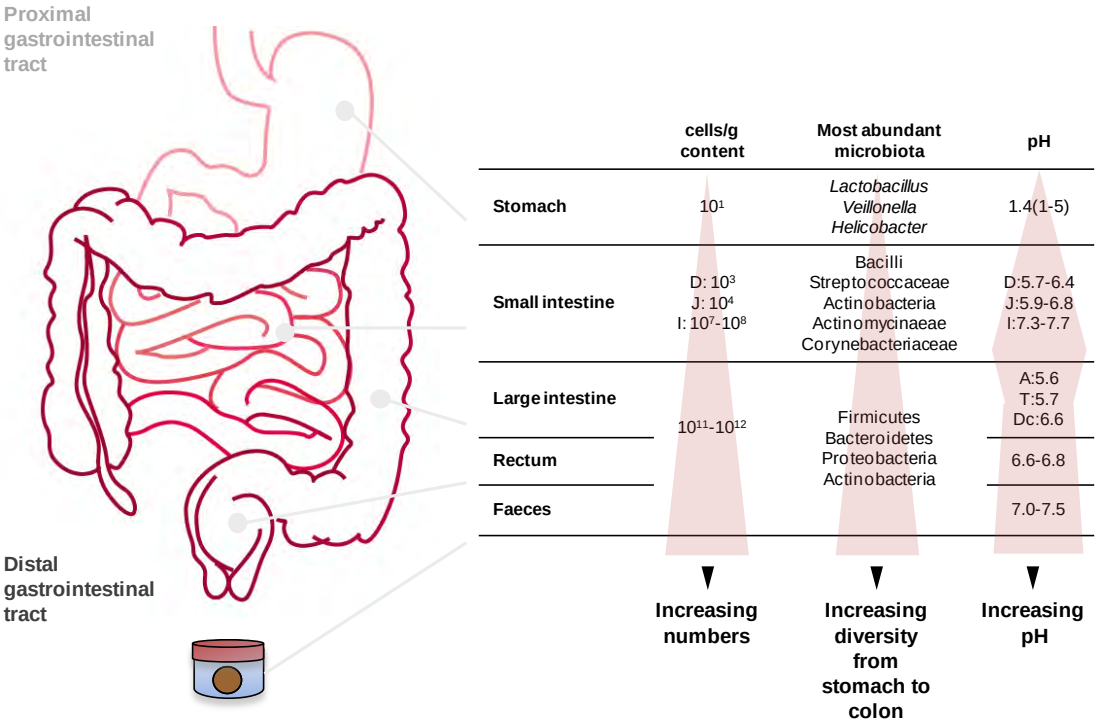


Figure 1.2. Variations in microbial numbers, composition and major factors that shape microbial community along the length of the gastrointestinal tract. D: duodenum, J: jejunum, I: ileum, A: ascending colon, T: transverse colon: Dc: descending colon, n.d: no data found. Adapted from [7, 8]).

Many factors shape the microbial diversity in the human gut [7, 9]. For instance chemical factors such as pH or redox potential, physical factors such as intestinal motility, and other factors such as differences in availability of water and nutrients, concentrations and bile salt contents, microbial competition, host pressure (mutualistic microorganisms are favoured), and the host's immune system determine the abundance and diversity of indigenous microorganisms in each location [7]. As a consequence, there are distinct microbial communities in each of the main regions of the gut. Microbial composition has been reported to be similar in the distal ileum, colon and rectum but differs from that found in the upper part of the gastrointestinal tract [10-13]. Facultative anaerobes such as *Streptococcus*, *Lactobacillus* and Enterobacteriaceae are predominant in the duodenum and jejunum. In contrast, strict

anaerobes such as Firmicutes and Bacteroidetes are more abundant in the distal ileum, colon and rectum (Figure 1.2).

In addition to longitudinal heterogeneity, there is also spatial organization. Consequently, the epithelial surface, the mucus layer and the lumen offer different environmental conditions or niches which support different microbial communities (Figure 1.3). The microbial composition of the lumen has been reported to differ from that found attached to the mucosa [10, 12], which in turn it is predominated by aerotolerant and asaccharolytic protein metabolizing bacteria [14]. This is due to the distribution of the tissue-associated mucus (which provides a nutrient source) and also to the radial oxygen gradient. While redox conditions in the colonic lumen are highly reducing (favouring the growth of strict anaerobes) higher oxygen tensions and lower redox potentials are found close to the mucosa [14]. Therefore, according to their position towards the mucus layer that covers the epithelium, colonic bacteria are classified as mucotrophes, feco-mucosal or mucophobes [15]. Also, the microbial community capable of colonizing food residues in the colon mainly consist of specialized species which act as primary degraders, whereas a more diverse community can be found in the liquid luminal phase [16].

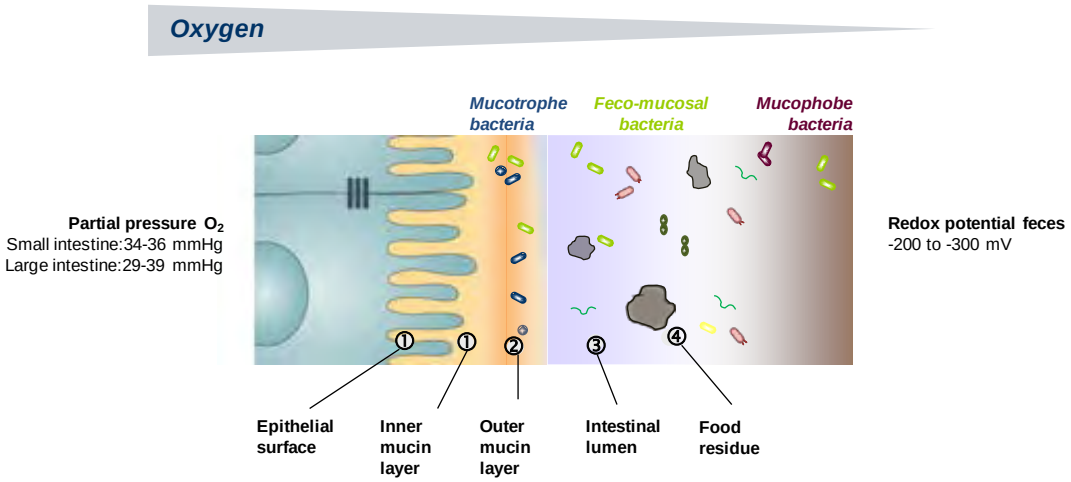


Figure 1.3. Major environmental factors, microbial communities and microbial microenvironments within the large intestine: ① epithelial surface and inner mucin layer (minimal colonization in the healthy state); ② diffuse mucin layer (specialist colonizers, such as *Akkermansia muciniphila*; ③ gut lumen-liquid phase (diverse microbial community); and ④ gut-lumen-substrate particles (specialized primary colonizers e.g. *Ruminococcus* spp.). Adapted from [7, 17].

Metagenomic studies have established that the human microbiome harbours between 1,000 and 1,500 bacterial species, but only around the twenty percent of these species have been cultured to date [10] which is attributable to their low relative abundance in the gut, rather than being inherently unculturable [18]. However, despite the high species richness, it has been evidenced that the gut microbiota is constituted by a relatively limited number of dominating bacterial phyla (Table 1.1).

Members of Eukarya (nine phylotypes belonging to Ascomycota), Archaea (*Methanobrevibacter smithii*, *Methanobrevibacter ruminantium* and *Methanosphaera stadtmanae*) and more than 1200 viral genotypes have also been detected within this complex community.

Table 1.1. Summary of the dominant (top 10) phylums, genus and species present in the human large intestine and their relative abundance.

Phylum*	Relative abundance (%)	Genus**	Relative abundance (%)	Species***	Relative abundance (%)
Firmicutes (mainly *Clostridium coccoides* (cluster XIVa) and *Clostridium leptum* (cluster IV))	57-82	*Faecalibacterium*	~5-8	*Ruminococcus bromii*	~0.01-10
Bacteroidetes	16-31	*Lachnospiraceae*	~2-5	*Dorea longicatena*	~0.35-8
Proteobacteria	<10	*Roseburia*	~1-4	*Faecalibacterium prausnitzii*	~0.28-8
Actinobacteria	<5	*Blautia*	~1	*Eubacterium halii*	~0.50-6
Fusobacteria	~1	*Coprococcus*	~1	*Ruminococcus torques*	~0.28-5
Verrucomicrobia	~0.2	*Ruminococcus*	~1-12	*Bacteroides caccae*	~0.03-18
Euryarchaeota	<0.1	*Bacteroides*	~5-18	*Bacteroides uniformis*	~1.51-15
Spirochaetes (two phylotypes)	<0.1	*Alistipes*	~1-4	*Parabacteroides merdae*	~0.02-15
Lentisphaerae (one phylotype)	<0.1	*Bifidobacterium*	~1-7	*Alistipes putredinis*	~0.08-10
Cyanobacteria (one phylotype)	<0.1	*Collinsella*	~1-3	*Bacteroides thetaiotaomicron*	~0.04-3

* data adapted from [10, 13, 19-23]; ** data based on the analysis of faecal samples of 163 individuals of whom: 146 European, 13 Japanese, 4 American; adapted from [24, 26] *** data based on faecal samples of 130 European adults; adapted from [25, 26]

Interestingly, only 160 of such species are shared among individuals and might constitute the human intestinal microbial phylogenetic core [26-29]. Thus, although there are some species in common among individuals, the entire intestinal microbiota is host-specific [10-12, 30-32].

Despite the diversity at the level of phylotypes or species, it has been established that most healthy adults usually have a relative stable climax microbial community in the colon at least for two years [30, 32], and it has been suggested that can be grouped into clusters characterized by a different bacterial ecosystem. Initially, three clusters (referred to as enterotypes) were defined. Each enterotype had dominant abundance of either *Bacteroides*, *Prevotella* or *Ruminococcus* [24, 33, 34]. Interestingly, enterotype clustering seemed to be mostly affected by long-term dietary habits and to be independent of nationality, age, sex, and body mass index [24, 33]. More recently, the existence of enterotypes has been questioned. Nevertheless, clustering of the gut microbial community based on relative abundance profiles of species is generally accepted [35].

1.1.3. Factors determining the gut microbiota composition

Gut microbiota is constituted by a dynamic community, in which the microorganisms undergo continuous selective pressure (for review see [36]). Therefore, the gut microbiota of an adult may suffer changes over time. Age, genetics, environment and diet are baseline factors that shape the gut microbiota composition of an individual. In addition, other factors

such as changes in diet, drugs intake, stress and host health can alter the gut composition of an individual over time.

It has been reported that there may be a genetic influence on microbiota composition since family members are found to have more similar gut microbiota than two unrelated individuals [37-40]. However, this can also be partially explained by shared environment [41]. Variations in gut microbiota composition between different cohorts of individuals from different ethnicities have also been observed [42]. For instance, major differences in the faecal microbiota between children from Africa and Europe have been recently reported based on analysis of amplified 16S rRNA gene sequences [43]. Interestingly, the *Bacteroides* phylum was more abundant in the African children, while *Firmicutes* were relatively higher in the European. Although the effects of genetics cannot be ruled out, the authors concluded that these differences are mostly attributable to differences in dietary habits.

Age- and gender-related differences in faecal microbiota have also been reported. In the elderly the number of enterobacteria is higher and the counts of Firmicutes lower than those found in working-age people (for review see [6]). The numbers of the *Bacteroides-Prevotella* group seem to be higher in males than in females [34].

Concerning the effects of drugs on the intestinal microbiota, one of the most perturbing and extensively used are antibiotics (for review see [44, 45]). It has been reported that antibiotics such as ciprofloxacin dramatically reduce the richness, diversity and evenness of bacterial taxons in adult faecal samples. Interestingly, it has also been shown that the composition of the community closely resembles the pre-treatment state one month after the end of treatment, although some taxa are not recoverable within six months [46]. However, the effects produced depend on the antibiotic used. For instance, β-lactams have been found to decrease the numbers of enterobacteria, enterococci, and anaerobic bacteria in several trials [45]. In contrast, antibiotics such as phenoxymethylpenicillin, metroinidazole, cefotaxime and several cephalosporins show only minor effects on disturbing the gut microbiota community [45]. It is noteworthy that, the imbalance in the microbial community induced by antibiotics can result in an increased susceptibility to pathogen colonization, for instance an overgrowth of organisms such as *Clostridium difficile* as a consequence of the elimination of other bacterial species that normally control its growth [45].

It is known that changes in diet can induce significant variations in gut microbiota composition (for review see [16]). The dietary intake of indigestible carbohydrates (e.g. resistant starch, non-starch polysaccharides, and prebiotics) affects the microbiota composition of the gut both in short-term dietary interventions and in response to long-term dietary intake. For instance, a significant increase in Ruminococcaceae (*Clostridium* cluster IV), *Eubacterium rectale*, *Roseburia* and *Oscillospira* has been observed in obese volunteers when

they switch from a diet enriched in non-starch polysaccharides to a resistant starch diet [16, 25]. Besides, the provision of diets with reduced carbohydrate intake in obese subjects has resulted in a significant decrease in the proportion and total numbers of bifidobacteria and butyrate-producing *Lachnospiraceae* related to *Roseburia* [47]. Also, fibre-restricted diets have been found to affect gut microbiota by decreasing the abundance of commensal beneficial bacteria such as *Faecalibacterium prausnitzii* and *Roseburia* spp. group [48]. Finally, a Western diet (enriched in total fat, animal protein, n-6 polyunsaturated fatty acids and refined sugars) has been demonstrated to cause a shift in the microbiota composition of mice comparable with what is observed in Crohn's disease (CD) patients, with an increase in the mucin-degrading bacterium *Ruminococcus torques*, and the group *Bacteroides/Prevotella* [49]. In line with these results, an association between the intake of fat, protein, and carbohydrate and the frequency of *Bacteroides*- and *Prevotella*- dominated microbiotas (enterotypes) has been found [33]. To be precise, faecal enterotypes rich in *Bacteroides* were associated with habitually high intakes of proteins and animal fat, whereas those rich in *Prevotella* were associated with higher carbohydrate intake [33].

Probiotics and prebiotics have also been widely used to modulate gut microbiota. Probiotics are living microorganisms that confer a health benefit on the host when administrated in adequate quantities [6, 50]. Many microbial strains which belong to the genera of *Lactobacillus*, *Bifidobacterium, Escherichia*, *Enterococcus*, *Bacillus* and *Saccharomyces* are commonly used in probiotic preparations [6, 50, 51]. In contrast, prebiotics are non-digestible food ingredients that, when eaten in adequate amounts, selectively modulate the growth and/or activity of particular microbial groups in the gut for their intended beneficial effect for the host. Inulin, trans-galacto-oligosaccharides, polydextrose, and resistant starches have been considered within this category of compounds as they can reach the colon and resist digestion of host enzymes, being therefore fermented mostly by the gut microbiota [52-54]. Inulin-derived prebiotics, for example, have been shown to result in significant increases in the representation of bifidobacteria and *F. prausnitzii* [55, 56].

The administration of a probiotic and a prebiotic simultaneously is known as synbiotic. Several trials have been performed in order to evidence the efficacy of probiotics, prebiotics or synbiontics in disease conditions such as atopic eczema, inflammatory bowel disease (IBD) and antibiotic associated diarrhoea [48, 53, 54, 57, 58]. However, responses were often individual-specific, probably due to differences in the initial composition of the microbiota before the intervention [16].

1.1.4. Functions of gut microbiota

Most of the microorganisms inhabiting the gut are either harmless or beneficial to the host. These commensal and symbiotic bacteria contribute to many important tasks for human health (for review see [59]). For instance, gut microbiota may influence host physiology such as glucose and lipid metabolism, brain function, etc. Beyond conferring additional metabolic activities to the host, gut microbiota has extensively been linked to host immunity development and also to play a role in maintain gut homeostasis.

Metabolic activities of gut microbiota

Despite the high host specificity, the overall metabolic activity of the intestinal microbiota seems to be similar amongst subjects [26, 36]. It has been estimated that the total number of microbial genes in these communities (referred to as microbiome) is between two million and four million, which represents approximately 150 times as many genes as the human genome [26]. Therefore, the microbiome embodies a vast metabolic potential which is greater than that possessed by the host, and which confers additional metabolic roles. Among the additional functionalities that gut microbiota confers on the host are those related with dietary product digestion and nutrient acquisition, essential vitamin production, participation in the metabolism of drugs and in the detoxification of toxic compounds.

Gut microbiota is involved in dietary nutrient release and complex polysaccharide break down, such as plant-derived pectin, cellulose, hemicellulose and resistant starches which the host is unable to digest. It has been demonstrated that the human gut microbiome is enriched in genes involved in starch and sucrose metabolism, as well as the metabolism of glucose, galactose, fructose, arabinose, mannose, and xylose [60]. The population levels of the main metabolic groups in the gut have been determined by culture-based and molecular approaches [61]. Starch degrading bacteria represent 10.1% of total viable counts, while mucin degraders accounted for 5.1%, and proteolytic bacteria for 1%. Among the fibre degrading population, xylanolytic bacteria accounted for 2.6%, and cellulolytic bacteria for 0.16% of total bacteria [61]. More recent studies have revealed that the minimal gut metagenome, which refers to bacterial functions involved in gut homeostasis and encoded across many species, relates to biodegradation of complex sugars and glycans harvested from the host diet and/or intestinal ligning [26]. This revealed the strong dependence of gut ecosystem on complex sugar degradation for its functioning.

In addition to dietary carbohydrate breakdown, gut microbiota has also been involved in the synthesis of aminoacids and essential vitamins for the host including thiamine, folic acid, nicotinic acid, pyridoxine, cyanocobalamin, biotin, and K vitamins [59]. Besides, gut microbiota is implicated in energy harvesting from diet through the production of short chain

fatty acids (SCFA) and by promoting the absorption of monosaccharides from the gut lumen because the induction of mucosal glucose transporters and hepatic lipogenesis is related with the intestinal microbiota [59, 62].

Finally, genes involved in the detoxification of xenobiotics have also been detected in the microbiome [60] and it has been reported that gut microbiota also induces the expression of major drug metabolizing enzymes for the host [63.].

Mutualistic relationship and homeostasis

Gut microbiota can be involved in host defence against infection in a direct manner, as the bacterial community competes for nutrients and epithelial binding sites with opportunistic pathogens (known as colonization resistance), or indirectly, by promoting host defences. For instance, it has been evidenced that the gut microbiota is involved in inducing the secretion of antimicrobial peptides and proteins of innate immunity e.g. angiogenin-4 [64]. In addition, the gut microbiota influences the development of the normal mucosal immune system through the induction of gut associate lymphoid tissue development, and the promotion of diversification of lymphoid populations and immunoglobulin genes [59]. Therefore, the healthy state in the human gut requires constant interactions and a delicate balance between the human host and the gut microbiota (for review see [64]), a situation known as homeostasis.

The gastrointestinal tract has evolved for maintaining the homeostasis with the commensal microorganisms that inhabit it without triggering the immune response. Thus, the immune system must be able to be tolerant to its indigenous microbiota, but at the same time has to be ready for an active response to pathogens. During microbial colonisation, the immune system matures, and the host develops a tolerance to commensal bacteria. Homeostasis is possible because systemic or localized immune responses against the commensal intestinal microbiota are prevented by a physical separation of bacteria and host cells, a down regulation of bacterial receptors and their ligands, as well as by a low ratio of pro-inflammatory/anti-inflammatory cytokines and the stimulation of protective molecules that mediate mucosal barrier function (induction of immune tolerance) [58]. In turn, also commensal bacteria have developed several mechanisms to crosstalk with the human cells (for review see [65]).

1.2. Dysbiosis and intestinal disorders

Since the human gut harbours a complex microbial ecosystem, which is capable of performing a variety of functions, it can be hypothesized that gut microbiota may be involved on many manners in regulating host health. For instance, the composition of the gut microbial community is assumed to be relevant to health because it determines the ratio of different microbial metabolites, the proportion of beneficial commensal organisms to potential pathogens, and the relative production of pro-inflammatory versus anti-inflammatory signals received by the immune system [16]. An emerging body of literature links imbalances in the gut microbial community, to multiple diseases. Dysbiosis, which refers to disturbances in the balance in the intestinal microbiota composition, has been pointed out to play a role in obesity [37, 38, 66], type 1 and type 2 diabetes [62, 67, 68], cancer development [7], allergies [64], fatty liver diseases [69], kidney disease [70], arthritis [71] and in neurological disorders like autism [72, 73]. However, most of the studies have focussed on depicting the association of gut microbiota with intestinal-related diseases, such as colorectal cancer [39, 74], irritable bowel syndrome [75-78], celiac disease [79, 80] and more extensively with inflammatory bowel diseases [27, 81-88]. A state of the art of the current knowledge about the role of gut microbiota in the intestinal diseases studied in this Thesis is detailed here.

1.2.1. Inflammatory bowel disease (IBD)

Inflammatory bowel disease (IBD) includes several chronic inflammatory disorders of the gut. Infectious colitis, ischemic colitis and radiation enterocolitis are IBDs of known aetiology, but there are still some IBDs of unknown causes. Among idiopathic IBDs, Crohn's disease (CD) and ulcerative colitis (UC) are the two major types [89].

Crohn's disease (CD)

Crohn's disease (CD) is a condition that mainly affects people in developed countries, with an incidence of 3-5 per 100,000 individuals annually [90]. CD can occur in all ages and genders, but the peak age of onset is around 20 years old. The distribution of prevalence among ages is bimodal, with a second peak of high prevalence existing for people between 50 and 70 years old [91].

Its common symptoms are pain, fever, bowel obstruction and bloody diarrhoea (for review see [58]). However, there are a great number of disease phenotypes, which can be categorized following the Montreal classification [92] depending on the age at the onset of the disease, the location of the inflammation, and its overall behaviour (Table 1.2).

Inflamed areas in CD patients are patchily distributed and may be found along the whole gastrointestinal tract (i.e., from the oropharynx to the anus) [89]. Inflammation can be transmural, thus affecting the whole intestinal wall from the mucosa to the serosa, which can lead to other complications such as fistulas, abscesses and stenosis.

Table 1.2. The Montreal classification of Crohn's disease [92].

Age of diagnose (A)	Characteristics
A1: younger than 16	Colonic localisation in most cases High family aggregation and genetic susceptibility
A2: 17-40 years old	Frequent and extensive inflammation, from upper gastrointestinal tract to colon
A3: older than 40	Colonic localisation in most cases
Localisation (L)	**Characteristics**
L1: ileal (I-CD)	30% CD patients Basic clinical manifestations: stenosis, nausea, vomiting, abdominal pain, loss of weight, and fever. Less aggressive diarrhoea than in colonic localisation.
L2: colonic (C-CD)	20% CD patients One or several affected areas between cecum and rectum, but mainly colon. Basic clinical manifestations: Abundant diarrhoea, bleeding, abdominal pain, and loss of weight. Correlates with perianal disease and extraintestinal manifestations.
L3: ileocolonic (IC-CD)	45% CD patients Localisation and clinical manifestations of L1 and L2
L4: upper gastrointestinal tract	5% CD patients Proximal ileum, jejunum, duodenum, stomach, oesophagus or oropharynx can be affected. Heterogeneous clinical manifestations depending on the exact localisation.
L4 localisation must be added to the above categories in case of additional upper GI tract involvement (e.g. L1+L4)	
Behaviour (B)	**Characteristics**
B1: Inflammatory (not stricturing-not penetrating)	Superficial ulcerations and inflammation Abdominal pain and diarrhoea
B2: Stricturing	Presence of stenosis and fibrosis Nausea, vomiting, pain and abdominal distension. Cases often refractory. Occasional surgical intervention. Low recurrence.
B3: Penetrating	Perforation. Often formation of fistulas and abscesses. Surgery necessary. High recurrence.
Perianal disease is a modulator of the above categories, which must be indicated with a p (e.g. B1p)	

Although CD aetiology has not been established yet, it has been evidenced that gut microbiota as well as environmental, genetic and immunological factors may be involved. Several hypothesis have been proposed (for review see [58]), but currently the favoured one is that the mucosal immune system shows an abnormal response towards luminal antigens such as commensal microbiota in genetically susceptible individuals [58, 93-95]. (For review see [96]).

The exact role of gut microbiota in CD development has been the focus of numerous studies in the last decade and is under constant research (for review see [93]). Nowadays there is a wide variety of clinical and experimental studies evidencing that gut microbiota is implicated in IBD [58, 94, 97]. For example, the presence of commensal microbiota has been demonstrated to be essential for the development of experimental colitis in several IBD animal models [97-100]. It has also been demonstrated that diversion of the faecal stream prevents recurrence in CD and the onset of pouchitis after surgery, and inflammation only develops after closure of the temporary ileostomy [101, 102]. Besides, some polymorphisms within the genes associated with innate immune responses to bacteria, such as NOD2 and autophagy-associated genes (e.g., *ATG16L1* and *IRGM*) have been linked to CD [103-105]. Culture-dependent as well as molecular techniques have showed that, in CD patients,

bacterial concentrations are increased while diversity is reduced [97, 106-109]. Finally, marked perturbations of the gut microbiota have been consistently reported in both faecal and mucosa-associated communities of CD patients [82-87, 93, 106, 110-114].

Although differences exist between different types of CD [87, 115, 116], a broad consensus has been reached on which species cause the bacterial imbalance observed in CD patients. Firmicutes phylum has been consistently reported to be less prevalent and abundant in CD patients [15, 81, 83-85, 111, 115]. Within this group, a reduction of *F. prausnitzii* is the most replicated species-specific finding so far, both in faeces and mucosa [85, 111, 114, 117, 118]. Also an increase of Enterobacteriaceae, especially *Escherichia coli,* has been systematically reported in CD patients in comparison to healthy subjects [15, 85, 115, 117, 119, 120]. In addition, depletion of other species as *Bifidobacterium adolescentis* and *Dialister invisus* has been pointed to characterise CD dysbiosis signature [83]. Also a qualitative and quantitative increase of the mucolytic species *Ruminococcus torques* and *Ruminococcus gnavus* has been found in CD [83, 85, 121]. At a subspecies level, the adherent-invasive *Escherichia coli* (AIEC) pathotype has been more frequently found in ileal-CD patients than in healthy controls [115, 122, 123] and has been systematically linked to many characteristics of CD pathogenesis, suggesting an aetiological role rather than being a consequence of inflammation [49]. However, to date, no single pathogenic bacterium has been conclusively shown as the cause of CD, and it stills remains to be elucidated if the dysbiosis is a cause, a consequence or may play a role in perpetuating the disease.

Ulcerative colitis (UC)

In contrast to CD, ulcerative colitis involves a more restricted area of the gut in an uninterrupted pattern, and is exclusively located in the colon and the rectum (for review see [124]). Unlike CD, inflammation is confined to the mucosa. Ulceration, edema and haemorrhage are also characteristic of UC. The incidence of UC has also been reported to be higher in developed than in developing countries, reaching values of 10 per 100,000 individuals annually [90].

Depending on the anatomical extension of the inflammation, three disease phenotypes can be defined (Table 1.3).

Although the precise aetiology of UC has remained indistinct, some genetic components are presumed to increase the disease susceptibility [125]. However, unlike CD, in UC a greater contribution of environmental factors is presumed as a significant lower monozygotic twin concordance rate has been reported in comparison to that of CD [88]. The possible implication of gut microbiota in UC initiation or development has been pointed out by several studies. Although studies have indicated that remission UC patients cannot be

discriminated from healthy controls [116, 126], many others support that UC patients harbour a rather instable microbial community that shows reduced diversity and richness in remission state [127, 128], and especially in active patients [126] and under clinical relapses [129, 130]. UC patients have increased numbers of mucosa-associated bacteria [131, 132] and alterations in the composition their gut microbiota community have been stated as well [88, 111, 128, 133-135] and in relation with their intestinal transcriptome, metabolome or proteome [127, 130, 136]. However, to reach a consensus is difficult since the initial studies on this field were performed in small cohorts and have used a wide variety of different methodologies and sampling protocols [106, 137, 138].

Table 1.3. Ulcerative colitis (UC) phenotypes [139].

UC phenotype	Characteristics
E1: Ulcerative proctitis	Inflammation is limited to the rectum. Generally mild intermittent rectal bleeding is the only symptom. The prevalence among UC patients is between 30-42%.
E2: Distal or left-sided UC	Involves the rectum, sigma and left colon. Symptoms include bloody diarrhoea, abdominal cramps, weight loss and left-sided abdominal pain. Prevalence among UC patients is 44-48%.
E3: Pancolitis or universal colitis	Refers to inflammation affecting the entire colon. Bloody diarrhoea, abdominal pain and cramps, weight loss, fatigue, and fever are symptoms of pancolitis. Patients with this disease phenotype have a higher probability of colectomy and colorectal cancer. Prevalence among UC patients is 9-17%.

The most reported observation has been a reduction of Firmicutes and Bacteroidetes, either at faecal or mucosal level [88, 106, 111, 134, 135, 137]. UC-dysbiosis has showed as well an increase of Proteobacteria, Actinobacteria, Fusobacteria, and Spirochaetes [128, 133, 140]. Besides, it has been shown that dysbiosis in UC patients is driven by different members of the microbiota than those responsible for CD-dysbiosis signature [110, 126]. The most comprehensive study on the field carried out by Machiels and collegues [88], where faecal samples from 127 UC patients were analyzed, has explicitly defined to species-level the UC dysbiosis. Precisely, there is a reduction in the abundance of known butyrate-producing species as *Roseburia hominis,* and the depletion of *F. prausnitzii,* previously pointed out only in UC active patients [117], has been corroborated. Although few studies separate patients according to activity status, in active UC patients an increase of sulphate-reducing bacteria and a reduction of some butyrate-producing bacteria have been described [117, 132, 141]. In addition, also an increase in concentrations of facultative anaerobes and proinflammatory bacteria has been reported [142, 143]. In line with these differences in functional groups, also changes in faecal organic acids composition, ammonia and indole have been observed in UC patients in comparison to healthy subjects [127].

Inflammatory bowel disease diagnosis

Although CD and UC are disorders that feature different localisation, histology and distribution of inflamed areas, sometimes they also share similar characteristics that hamper a clear classification. In addition, IBDs feature a relapsing course, with disease-active periods

or flares, which alternate with inactive episodes of remission. Therefore, as clinical manifestations are unstable during disease course, a long monitoring period (at least five years) is necessary to accurately classify the disease phenotype [144]. Currently, clinical features in conjunction with image-based tools and histology are necessary for accurate IBD diagnosis. **Endoscopic techniques** allow an overall description of the amount and characteristics of the lesions as well as the disease location. Besides, biopsy sampling is possible during endoscopy. **Radiology** provides important additional data about the behaviour of the disease (e.g. presence of fistulas), which makes it useful for phenotype classification. Many **histological features** are used for IBD diagnosis [145], with the presence of granulomas being a key characteristic of CD [146]. Several **biological markers** are used in combination with the previously mentioned diagnostic techniques and can be assessed in blood or in faeces [147]:

- Perinuclear anti-neutrophilic cytoplasmatic antibodies (pANCAs) and anti-*Saccharomyces cerevisiae* antibodies (ASCAs)
- C-reactive protein quantification (CRP)
- Globular sedimentation rate determination
- Faecal calprotectin and lactoferrin concentrations

None of these biological markers is pathognomonic for either UC or CD, thus they are used as supplement to endoscopy. However in some cases they can be useful for CD and UC differential diagnosis. For instance, ASCAs are more characteristic of CD than UC (50-60% of CD patients show positive results, whereas only 10% of UC patients are ASCAs-positive). In contrast, 70% of UC patients are pANCAs-positive whereas only between 10-40% of CD patients have positive result for this serological biomarker. In addition, some of them can be useful to classify the disease phenotype and estimate disease course (e.g. CRP quantification), as well as for prognostic purposes such as predicting the response to treatment. New serological markers have been associated with IBDs (for review on this subject see [148]).

Despite the technical difficulties involved in studying the gut microbiota, the differences observed between IBD and healthy subjects set the rationale to implement *in vitro* diagnostics of the human gut microbiota (for review in this subject see [149]). Recently, some studies have started to explore the potential applicability of dysbiosis of the gut microbiota and/or its metabolites for IBD diagnosis as a novel strategy which may support disease diagnostics or prognostics. Swidsinski and colleagues have reported that active CD and UC could be diagnosed taking into account *F. prausnitzii* abundance in conjunction with faecal leucocytes counts [114]. Recently new phylogenetic specificities of CD microbiota have been highlighted, and their usefulness to discriminate CD patients with ileal involvement has been suggested [86]. Finally, a study based on the analysis of faecal samples of a Japanese cohort

has demonstrated the feasibility of using the faecal microbiota profile as a predictive marker for CD activity [150]. Concerning metabolite profiling, Williams *et al.* (2009) demonstrated that CD and UC patients differ in urinary metabolites related to gut microbial metabolism, and identified potential biomarkers to specifically distinguish these two IBD from each other and from control individuals [151]. Therefore, although accumulating evidence suggests that gut microbiota may be a useful source of additional information to assist in IBD diagnostics, and some studies are attempting to implement bacterial indicators or related metabolites with this goal, to date there are still no tools set up which may be of assistance in IBD subtypes differential diagnose. In addition, there is also a lack of comprehensive studies showing how patients' clinical data correlate with changes in the abundance of these bacterial indicators, and how the different therapies may affect the abundance of these species, which is also an important issue to consider in order to fully implement bacterial biomarkers as a supporting diagnostic tool.

Inflammatory bowel disease treatment

IBD management depends on the disease location, severity and activity. Common therapies currently used are anti-inflammatory chemicals derived from salicylic acid (i.e. mesalazine, and sulfasalazine), corticosteroids (i.e. prednisone, methyl-prednisone and budesonide), antibiotics (i.e. metronidazole and ciprofloxacin), immunosupressors (i.e. azathioprine and mercaptopurine), antimetabolite and antifolate methotrexate, and the so called "biological " drugs consisting of antibodies against tumour necrosis factor α (TNF-α), such as infliximab and adalimumab. Intestinal resection is also indicated in those patients with fulminant or fistulising CD and for those patients unresponsive to any of the previously mentioned medication (refractory cases). For review see reference [152]. More recently, persistence of unmet therapeutic needs in CD patients with refractory disease has raised interest in innovative cellular immunoregulatory and regenerative medicines including autologous hematopoietic stem cell transplant [153-155]. Also a growing body of literature supports the emerging concept that suggests that probiotics or prebiotics may have therapeutic effects in IBD through balancing the dysbiosis [156-159]. For instance, studies in animal models have pointed out that some species of the gut microbiota such as *Bacteroides fragilis* and *F. prausnitzii* are able to produce molecules that prevent colitis or with anti-inflammatory effects respectively [118, 160], which shed new light on the future use of gut microbiota as therapeutics in this intestinal disorders.

1.2.2. Irritable bowel syndrome (IBS)

Irritable bowel syndrome (IBS) is a poorly understood condition which usually starts in early adult life, and affects between 9% and 22% of the population in the United States and Europe [77, 161].

Patients suffering from IBS may have abdominal pain and/or discomfort, bloating, excessive flatulence, and bowel disturbances [162]. Depending on clinical symptoms, IBS can be subdivided in three subtypes: diarrhoea-predominant IBS (IBS-D), constipation-predominant IBS (IBS-C) and both (IBS-M).

Several diagnostic criteria (Kruis, Mangin, Rome) have been used to distinguish IBS patients form those with organic bowel disease in daily clinical practice [163, 164], but the most recent and currently used is the Rome III criteria (available at <http://www.romecriteria.org/criteria/>). According to these criteria, developed to classify the functional gastrointestinal disorders, a subject is eligible to be suffering IBS if there is a recurrence of abdominal pain or discomfort at least 3 days/month in the last 3 months, associated with two or more of the following conditions:

- Improvement with defecation
- Onset associated with a change in frequency of stool
- Onset associated with a change in form (appearance) of stool

This diagnostics criterion should be fulfilled for the last 2 months with symptoms initiation at least 6 months prior to diagnosis.

IBS is a disorder previously thought to be exclusively psychosomatic [165, 166], as clinical symptoms of gastrointestinal dysfunction are present, but clear endoscopic and histological evidence is absent. The exact aetiology and pathophysiology of IBS remains to be elucidated, although several hypotheses have been proposed. Influence of possible alterations in the central and enteric nervous systems, as well as impaired permeability triggered by a luminal antigen, and altered levels of gastrointestinal neuropeptides and hormones have been indicated as possible causes or contributing factors of the disease [75, 77, 163, 166-169]. Furthermore, abnormal intestinal motility, as well as genetic, environmental, and physiological factors (e.g. anxiety, interpersonal sensitivity, depression, hostility, etc.), may also play important roles in the development of IBS [166, 170].

In addition, recent research findings have revealed that IBS patients feature alterations in colonic fermentation, and that the gut microbiota may be relevant for the disease pathogenesis [169, 171]. For instance, an increased production of hydrogen gas has been reported in IBS-D patients, whilst methane gas is produced in larger amounts in those patients suffering of IBS-C [172]. The predominant gut microbiota is highly instable over time in IBS subjects in comparison to controls [173, 174]. Cultured-based approaches have evidenced that there is an increased quantity of aerobic bacteria and *Lactobacillus* in IBS with respect to healthy subjects [173]. Studies based on molecular methods have further defined the differences in the microbiota of IBS patients with respect to that of controls. A significant increase of Bacteroidetes, Synergistetes and *Clostridium coccoides* subgroup and a

reduced abundance of Actinobacteria such as the genus *Collinsella,* Bacilli, Flavobacteria, Epsilonproteobacteria and some members of the phyla Firmicutes in IBS than controls [76, 166, 169]. Besides, changes in the relative abundance of Firmicutes and Bacteroidetes have been observed in IBS patients [175]. Precisely the Firmicutes were twice the counts of Bacteroidetes, partially due to an increase in the quantity of *Ruminococcus, Clostridium,* and *Dorea* species with a concomitant decrease in *Bifidobacterium* and *Faecalibacterium* species [176]. Although reports describing differences between IBS disease phenotypes are scarce, IBS-D and IBS-C appear to have distinct microbial populations. Lower amounts of *Lactobacillus* spp. have been observed in IBD-D patients whereas increased amounts of *Veillonella* spp. have been reported in IBS-C patients [76], and also differences in dominant subgroups of clostridia have been observed between IBS-C and IBS-D subjects [174]. The variety of disease subtypes currently included within IBS, in addition to the variety of techniques and samples used by the different studies, may be hampering to reach a consensus on IBS-dysbiosis signature. For review on this subject see [77, 78, 166].

Interestingly, recent studies have also shown that IBS is associated with low grade intestinal inflammation resulting from an activated immune system, in response to a normal or abnormal gut microbiota [166, 175, 177]. Thus, since IBS-D and CD share some similarities, it has been hypothesised that IBS is an inflammatory disease that shares common pathogenic features with CD, but has a milder phenotype [169]. Since some of the symptoms are similar to those observed in IBD, it is a challenging task for clinicians to accurately diagnose these two different intestinal conditions, particularly in the early stages of the disease.

1.2.3. Colorectal cancer (CRC)

Colorectal cancer (CRC) is the fourth most common cause of cancer-related mortality in the world and affects 6% of individuals by the age of 75, being the incidence much greater in developed than developing countries [90].

Development of colorectal cancer has been associated with many factors such as age, diet and genetic predisposition. However, increasing evidence suggests that the gut microbiota may play an important role in the pathogenesis of CRC (for review see [178]). To date, two mechanisms through which gut microbiota may be associated to CRC have been proposed. On the one hand, gut microbiota may promote chronic inflammation which in turn can lead to tumour formation. In this sense, members of the Enterobacteriaceae family as well as *Clostridium difficile* have been proposed as candidates to promote CRC development through the induction of chronic inflammation of the gut [179, 180]. Besides Uronis and co-workers, 2009 [74] demonstrated in a murine model the existence of a colitis-associated cancer in which microbial recognition system promotes the tumour development.

On the other hand, it has been evidenced that as a result of dietary compounds metabolism by the gut microbiota, some carcinogenic compounds such as ethanol, heterocyclic amines, hydrogen sulphide, and oxide radicals are formed [181-185]. Besides, some members of the genus *Clostridium* spp. and *Eubacterium* spp. have been reported to be able to transform deconjugated primary bile acids via 7-α-dehydroxylation into deoxycholic and lithocholic acids, which are potential carcinogens [90, 186]. In contrast, it is thought that the production of butyrate by gut commensals may provide some protection against CRC. In line with these observations, a decrease in the main butyrate-producers has been reported in faecal samples of CRC patients in comparison to controls [187]. These results are supported by a recent study comparing a cohort of African Americans with native Africans which has evidenced that CRC risk is influenced by the balance between microbial production of health-promoting metabolites such as butyrate, and potentially carcinogenic metabolites such as secondary bile acids [188].

Given the evidences on the putative **role of gut microbiota** in promoting CRC, in the last years some studies have faced the question to elucidate if there is a CRC-specific dysbiosis, or if there is any particular species which can be associated to CRC development. Among the most replicated findings there the association of members of the genus *Bacteroides* with CRC [189-191] although other genus such as *Streptococcus*, *Enterococcus*, and *Escherichia* have also been associated by some studies [182, 185, 192, 193]. A pyrosequencing study of Wu and colleagues [189] showed that 17 phylotypes closely related to *Bacteroides* were enriched in the gut microbiota of CRC patients as well as some potentially pathogenic bacteria as *Fusobacterium* and *Campylobacter*. In contrast, nine OTU, represented by *Faecalibacterium* and *Roseburia* were significantly less abundant. Therefore accumulating evidence suggests that the dysbiosis signature in CRC patients is characterised by a decrease in butyrate-producing bacteria, and an enrichment of potential pathogens. At the moment, it remains unknown if these changes in the gut microbiota of CRC patients are due to the disease or not, but mounting evidences link the microbiome with CRC pathogenesis, and the most plausible hypothesis is that several species with common pathways may be playing a role to promote tumorigenesis [178].

1.3. *Faecalibacterium prausnitzii,* a commensal bacterium of the human gut microbiota

In view that main members of gut microbiota may play a crucial role in disease onset and that *Faecalibacterium prausnitzii* has been reported to be part of the dysbiosis observed in various intestinal disorders, particularly in IBD, in this work it was decided to focus in studying this species.

1.3.1. Phylogeny and genome information

Faecalibacterium prausnitzii was initially named by Moore and Holdeman in 1973 as *Fusobacterium prausnitzii* [194, 195]. However, further analysis based on the 16S rRNA gene sequence demonstrated that strains classified phenotypically as *Fusobacterium prausnitzii* were not phylogenetically related to true *Fusobacterium* species [19, 196]. Duncan and co-workers reviewed this species' phenotypical traits and phylogeny and established, on the basis of the 16S rRNA gene sequence and the average GC content in its genome (47-57%), that the former *Fusobacterium prausnitzii* were more closely related to members of *Clostridium* cluster IV (the *Clostridium leptum* group), and the new genus *Faecalibacterium* was created [197]. Currently, *F. prausnitzii* is the only faecalibacteria species isolated, and is one of the main representatives of the Firmicutes phylum, *Clostridium* class, and Ruminococcaceae family found in the human gut [198, 199].

According to data released in the National Center for Biotechnology Information in July 2015 (accessible in http://www.ncbi.nlm.nih.gov/) there are currently five *F. prausnitzii* strains whose complete genome sequence is available. Two of these genomes, corresponding to strains SL3/3 and L2/6 are completely sequenced and annotated, and the former is considered the representative genome of the species. Three others are in progress (strains A2-165, M21/2 and KLE1255) but the annotations are still incomplete. Sequenced *F. prausnitzii* strains appear to lack plasmids and have circular 2.93 to 3.32 Mb chromosomes, which encode from 2,741 to 3,493 predicted proteins.

1.3.2. Phenotypic characteristics and isolation methods

F. prausnitzii is a low-GC, non-spore-forming rod-shaped bacterium of approximately 2 μm in length (Figure 1.4A and 1.4B). Despite being classified within the Gram-positive Firmicutes phylum, all *F. prausnitzii* isolates tested so far have responded negatively to Gram staining [197]. Colonies on M2GSC or YCFA media [197, 200] are <1 mm in diameter, and have an opaque to translucent appearance (Figure 1.4C).

This non-motile, non-flagellated bacterium produces butyrate, D-lactate and formate, and utilizes acetate during glucose fermentation. All *F. prausnitzii* isolates can grow on fructose, fructo-oligosaccharide, starch and inulin; whereas none can utilize arabinose, melibiose, raffinose, rhamnose, ribose and xylose. The isolates differed in their ability to ferment cellobiose, maltose and melezitose [197]. However, further characterisation of a larger number of strains, isolated from a wider range of subjects (*i.e.* from different age, and/or ethnic groups, with different diets, and/or suffering different intestinal disorders) would be of interest in order to improve understanding of nutritional requirements and factors crucial for the survival of this bacterium in the gut, intra-species phenotypical

diversity and to give insights into novel methods or strategies to improve its isolation and cultivation *in vitro*.

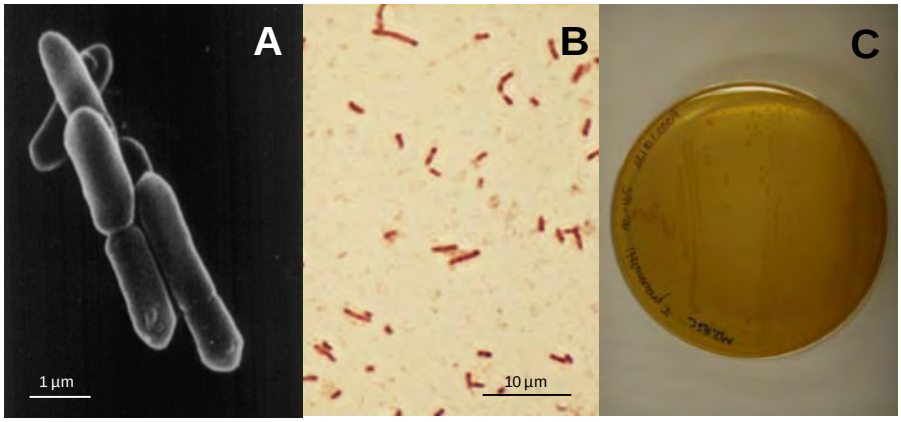


Figure 1.4. *F. prausnitzii* A2-165 cellular and colonial morphology. (A) Scanning electron micrograph obtained from [197]. (B) Optical microscope photography of *F. prausnitzii* A2-165 Gram staining. (C) *F. prausnitzii* grown on M2GSC medium [200].

F. prausnitzii is a bacterium that has been difficult to culture since first isolated by Prausnitz from pus from a case of pleural empyema in 1922 [195, 201]. Few attempts have been made to isolate and phenotypically characterise new strains. By the time to start this Thesis, apart from the initial work of Prausnitz, there are only three other studies where isolation of *F. prausnitzii* from human samples has been reported [195, 202, 203]. All the isolates described have been obtained through protocols of massive isolation of gut bacteria from faecal samples of healthy individuals. No strategy of enrichment or a selective isolation method for this species has been established so far. The difficulty of culturing *F. prausnitzii* through conventional methods is on the one hand due to the fact that its metabolic requirements are not well know yet, and on the other because it is an extremely oxygen sensitive bacterium. However, it has been recently shown that it can be grown in micro-aerobic conditions when flavins and cysteine or glutathione are present in the culture medium [204], an interesting ability that merits further investigation in order to assess if the addition of these compounds in the growth medium may ease its culture manipulation in aerobic conditions. If so, this could be the starting point for developing new isolation methods based on more user-friendly techniques. Most recent research supports this suggestion, because it has been shown that up to 60% of *F. prausnitzii* can survive 24h exposed to air when formulated with cysteine, riboflavin and inulin [205].

1.3.3. Molecular methods to study *F. prausnitzii* prevalence and abundance

Given the difficulty in culturing *F. prausnitzii*, the study of this species through traditional microbiological methods is extremely time-consuming and laborious. Therefore,

with the growing interest in this species, in the last few years several molecular methods have been developed to detect and/or quantify this bacterium in gut samples. Table 1.4 summarizes the main features of the methods described to date.

Table 1.4. Summary of molecular methods for *F. prausnitzii* detection and/or quantification, specificity and sample types in which it has been applied successfully.

Method *	Name*	Target gene*	Sequence (5'→3')*	*F. prausnitzii* strains validation	Samples used	Ref
cnPCR	FPR-1 FPR-2	16SrRNA	F: AGATGGCCTCGCGTCCGA R: CCGAAGACCTTCTTCCTCC	ATCC27768, ATCC27766	faeces, biopsies	[196]
	Fp.ID.F2 Fp.ID.R2	nucleotidyl transferase butyryl-CoA transferase	F: GTGACCGGATCGAACGACC R: TCCAGGTCATGTGGGCAGC	A2-165, M21/2.	faeces	[206]
FISH	Fprau645	16S rRNA	CCTCTGCACTACTCAAGAAAAAC	A2-165	faeces, biopsies	[198]
	Fprau655	16S rRNA	CGCCTACCTCTGCACTAC	A2-165, L2-6	faeces, biopsies	[207]
qPCR (SYBR)	nn nn	16S rRNA	F: CCCTTCAGTGCCGCAGT R: GTCGCAGGATGTCAAGAC	*in silico* (FASTA3 and Probe Match RDP)*	faeces, biopsies	[208]
	Fprau 07 Fprau 02	16S rRNA	F:CCATGAATTGCCTTCAAAACTGTT R: GAGCCTCAGCGTCAGTTGGT	A2-165, L2-6	faeces, tissue	[117]
	FPR-1 FPR-2	16SrRNA	F: AGATGGCCTCGCGTCCGA R: CCGAAGACCTTCTTCCTCC	ATCC27768, ATCC27766	faeces, biopsies	[196]
	FPR-2F FPrau645R	16S rRNA	F:GGAGGAAGAAGGTCTTCGG R:AATTCCGCCTACCTCTGCACT	A2-165	faeces, luminal aspirate	[55, 196]
	Fprau223F Fprau420R	16S rRNA	F: GATGGCCTCGCGTCCGATTAG R: CCGAAGACCTTCTTCCTCC	A2-165	faeces, biopsies	[196, 209]
qPCR (TaqMan)	Fprau161-177F Fprau180-196P Fprau215-199R	16S rRNA	F: CCCGGCATCGGGTAGAG P:AAAAGGAGCAATCCGCT R:GGACGCGAGGCCATCTC	A2-165	faeces	[83]

*cnPCR, conventional polymerase chain reaction; qPCR, quantitative PCR; FISH, fluorescent in situ hybridization; nn, no named; F, forward primer; R, reverse primer; P, probe; RDP, Ribosomal Database Project

Conventional PCR (cnPCR)-based methods offer an easy and cheap strategy for detecting *F. prausnitzii* in a wide range of samples. Two primer sets, one of which targets functional genes and differentiates two subgroups within *F. prausnitzii,* have been set up to date. However, systems for quantitative real time PCR (qPCR) can also be used in a cnPCR fashion (i.e. primers FPR-1 and FPR-2).

Fluorescent *in situ* hybridisation (FISH) and qPCR are two well established techniques for detection and quantification of bacteria in complex microbial communities. Two FISH probes have been described to date for *F. prausnitzii* [198, 207]. Although targeting similar regions of the 16S rRNA gene, the oligonucleotide designed by Suau *et al.*, (2001) has been more extensively used [198]. FISH is an efficient method for the direct quantification of bacteria and also, as it is based on microscopy, it evidences the target bacteria organization within the microbial community. However, it is time consuming, and in complex microbial communities might not be suitable due to the formation of dense microbial clusters which may compromise probe hybridization. In contrast, qPCR is a rapid and reliable method that overcomes FISH's main handicaps, but an efficient method for DNA extraction from the sample is required in order to minimise quantification bias, and no information on community structure is obtained. Several primer sets targeting *F. prausnitzii* 16S rRNA gene have been reported based on SYBR Green as fluorescence reporter dye. SYBR Green binds to double-stranded DNA, thus it can be easily implemented in any previously described PCR primer set (i.e. FPR-1 and FPR-2). However, as the method to report fluorescence is not sequence-specific, an additional step consisting of a melting curve

has to be performed in order to verify that the fluorescence recorded actually has originated from amplification of the desired target sequence. A recent study has compared the specificity of three *F. prausnitzii* –specific PCR primers pairs and has proved that to date, all the three primers can detect *F. prausnitzii* and *Subdoligranulum* spp. However, the specificity of the FPR-2/Fprau645R was shown to be better than FPR-1/FPR-2 [210].

Quantitative PCR assays based on hydrolysis probes (*i.e.* TaqMan) include, in addition to the primer set, a probe that reports fluorescence when there is successful amplification of the target sequence. The inclusion of this third oligonucleotide gives extra specificity to the assay and may help to overcome the specificity issues reported for the previously described primer sets. In addition, probes can be labelled with different fluorescent dyes, allowing qPCR-multiplexing. However, at the time of starting this Thesis, no hydrolysis probes assays had been reported for the quantification of *F. prausnitzii,* and to our knowledge only one assay has been described to date [83].

Although all the molecular methods described so far are used to evaluate clinical samples, none have included inhibition tests, mandatory for confirming negative results. Furthermore, although the validation of most of the methods included *in silico* tests, the *in vitro* specificity tests were restricted to only a few *F. prausnitzii* strains and only in some cases an exclusivity test including other gut species was performed. The design and implementation of a molecular tool to quantify *F. prausnitzii* meeting all the criteria required by laboratories of clinical diagnostic is of interest. Apart from providing a reference molecular tool to study this species abundance in different gut conditions, it could potentially be used as a biomarker, which in turn would be a source of additional information to assist in clinical practice.

1.3.4. Metabolism and role in the gut

F. prausnitzii ability to switch metabolism between strictly anaerobic growth (most likely to be found in the gut lumen) and micro-aerobic conditions (likely to be found close to the intestinal mucosa) may explain its ubiquity in the colon. It has been found associated with the intestinal mucosa forming two types of adherence: coat- or string-like adhesion and patchy adhesion [15]. A recent study based on a laser-captured micro-dissection technique has evidenced that *F. prausnitzii* is one of the most abundant species (20-50%) among mucosa-associated microbiota [211], and it has also been found in faecal samples and as part of the bacterial biofilms colonizing food residue [212]. *F. prausnitzii* might therefore be considered as a feco-mucosal bacterium.

The role of *F. prausnitzii* in the gut is starting to become clear. It has been consistently reported as one of the main butyrate producers found in the intestine, as it can generate

more than 10 mM of this short chain fatty acid in batch culture [202]. Butyrate plays a crucial role in gut physiology and host wellbeing. It is the main energy source for the colonocytes and it has protective properties against colorectal cancer and inflammatory bowel diseases [213, 214]. Butyrate can reduce intestinal mucosa inflammation through inhibiting NF-κB transcription factor activation [215], upregulating PPARγ [216] and inhibiting interferon gamma (IFN-γ) [217].

Additional anti-inflammatory properties have been elucidated in *F. prausnitzii,* as summarized in Figure 1.5. The *in vitro* stimulation of peripheral blood mononuclear cells by *F. prausnitzii* induced a tolerogenic cytokine profile with very low secretion of pro-inflammatory cytokines like IL-12 and IFN-γ, and an elevated secretion of the anti-inflammatory cytokine IL-10 [118, 218]. *F. prausnitzii* cells or their cell-free supernatant clearly reduced the severity of acute [118], chronic [219] and low grade [220] chemical-induced inflammation in murine models. These anti-inflammatory effects were partly associated with secreted metabolites capable of blocking NF-κB activation, IL-8 production [118] and upregulation of regulatory T cells production [218]. Recently seven peptides that derive from a single microbial anti-inflammatory molecule, a15 kDa protein, have been identified in *F. prausnitzii* cultures supernatant, and their capability to block NF-κB pathway has been demonstrated [221]).

F. prausnitzii supernatant has also been shown to attenuate the severity of inflammation in mice by affecting permeability thus enhancing the intestinal barrier function [220, 222]. The mechanism by which *F. prausnitzii* ameliorates permeability seems to be related with expression of certain tight junction proteins, but not with an enhancement of claudin expression [222]. Besides, a recent study performed using a gnotobiotic model has shown that *F. prausnitzii* could also influence gut physiology through mucus pathway and the production of mucus O-glycans, and may help to maintain suitable proportions of different cell types of secretory linage in the intestinal epithelium [223]. Finally, a restoration of serotonin (a key neurotransmitter in the gastrointestinal tract that affects motility [224]) level to normal has been evidenced in murine models treated with either *F. prausnitzii* or its supernatant [220].

Li *et al.* (2008) demonstrated that *F. prausnitzii* is an active member of the microbiome that influences numerous host pathways, as its population variations were associated with the modulation of urinary metabolites part of different host pathways such as tyrosine metabolism [226]. This link has been corroborated in faecal samples of healthy subjects [227]. More recently, the protective effect of *F. prausnitzii* has been linked with metabolites of the gastrointestinal tract such as the anti-inflammatory shikimic related to the salicylic acid pathway [228].

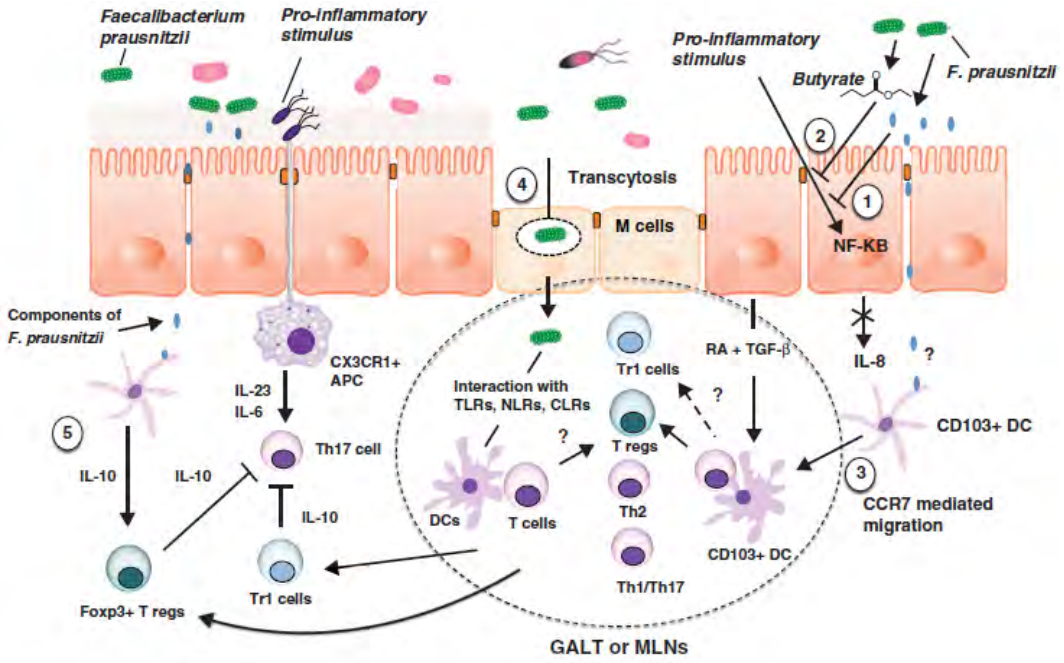


Figure 1.5. Schematic view of the suggested anti-inflammatory mechanisms of *F. prausnitzii* (adapted from [225]).① NF-κB activation induced by a pro-inflammatory stimulus may be blocked by some components from the supernatant of *F. prausnitzii* [118]. ② Butyrate produced by *F. prausnitzii* prevents NF-κB activation at mucosal level. ③ *F. prausnitzii* components may stimulate the migration of CD103+ dendritic cells (DCs) from lamina propria to gut-associated lymphoid tissue (GALT) or mesenteric lymph nodes (MLN), which will induce Tregs. ④ M cells transcytosis of *F. prausnitzii* in organized lymphoid structures might also induce T regs ⑤ The capacity of *F. prausnitzii* to induce high amounts of IL-10 in antigen presenting cells may enhance the suppressive activity of Foxp3+ Tregs and block Th17 cells induced by pro-inflammatory stimuli.

Finally, another point supporting that *F. prausnitzii* is a functionally important member of the microbiota is the fact that it is widely distributed over the animal kingdom [225]. To date, it has been reported that *F. prausnitzii* can also be found among the gastrointestinal microbiota of pigs [229], mice [211, 230], and calves [231] as well as poultry and the insect cockroach [232, 233], thus evidencing that it is not an exclusive gut symbiont.

1.3.5. *F. prausnitzii* in healthy individuals

It has been noted that *F. prausnitzii* is not detectable in faecal samples of babies less than six months of age [234]. Children of one to two years of age already have a significant amount of *F. prausnitzii*, and some of the current isolates available of this species have been recovered from an infant two years old [202]. No significant difference between adults and young children in the average proportion of F. prausnitzii has been found [4].

Although it is not well established when *F. prausnitzii* colonization of the gut might take place, it seems likely that it may be after weaning and the introduction of solid food,

when it has been reported that an increase in the abundance of Firmicutes takes place in infants [5]. A recent study has evidenced that it was not possible to obtain successful implantation of *F. prausnitzii* in gnotobiotic rats, and that it was needed the presence of *B. thetaiotaomicron*, which supports the hypothesis that *F. prausnitzii* is not a primarily colonizer of the gut. *F. prausnitzii* co-occurrence with several members of Bacteroidetes has been demonstrated in an *in silico* study performed by Lozupone and collegues, 2012 [235]. Therefore, it is plausible that *F. prausnitzii* would arrive to the gastrointestinal tract late after birth, once it has been colonized by bacteria such as *B. thetaiotaomicron,* which is less demanding concerning physicochemical conditions of the gut which in turn is a crucial factor governing strict anaerobes gut colonization and maintenance.

Once established in the gut, *F. prausnitzii* becomes a key representative of the intestinal microbial community. It is one of the three most abundant species found in adult human faeces by anaerobic cultivation [191] and by 16S rRNA-based molecular studies [19, 25, 29, 198, 207]. The average abundance in adult faecal samples ranges from 10^8 to 10^{11} cells/g of faeces [83, 196, 208, 236]. Currently it is accepted that *F. prausnitzii* represents between the 2 and the 15% of the total Bacteria sequences, and about the 8% of the Firmicutes sequences recovered in diversity studies of the human gut based on the 16S rRNA gene from faecal samples [10, 25, 198, 207, 237]. Similar proportions have been found when mucosa-associated communities have been analyzed in healthy individuals [15, 115], although a recent study based on laser-captured micro-dissection technique has evidenced that this values can increase to around 20-50% in some individuals [115, 211].

In the frail elderly, *F. prausnitzii* abundance is reduced [238, 239]. It has been hypothesised that slow bowel transit which also impacts on metabolic products present in the colon as well as colonic pH may be implicated in this fact [6].

The causes of the observed variations in abundance between individuals and ages are yet to be elucidated. On the one hand, *F. prausnitzii* abundance can be influenced by other members of the gut microbiota. It has been shown that *F. prausnitzii* co-occurs with Bacteroidetes or Clostridium cluster XIVa species (*C. coccoides* group) [235] and is also positively stimulated *in vitro* by the addition of exopolysaccharides from *Bifidobacterium pseudocatenulatum* in faecal cultures from some individuals [240].Thus, changes in some of these species may influence in turn *F. prausnitzii* numbers and may explain the observed intra-individual variations. Besides, although none of the enterotypes described is characterised by a *F. prausnitzii* dominant population, the *Prevotella* and *Ruminococcus* enterotypes have been reported to have respectively a positive and a negative co-occurrence of Ruminococcaceae members (including *F. prausnitzii*) [24, 33]. Thus, it can be hypothesised

that the stable community (enterotypes) of an individual can also explain, at least partially, the differences in *F. prausnitzii* abundance within healthy individuals.

On the other hand, differences in *F. prausnitzii* abundance can also be influenced by host factors such as gender. A recent study has demonstrated that Chinese men possess significantly lower numbers of faecal *F. prausnitzii* than women, suggesting that it may derive from gender difference in hormone and/or gut physiology. Finally, subjects' disease status or diet can also modulate *F. prausnitzii* numbers in the gut (discussed below, sections 1.3.6 and 1.3.7).

1.3.6. Abundance in different intestinal disorders

Several studies have shown that *F. prausnitzii* prevalence and abundance is reduced under certain intestinal disorders, particularly IBD. This raised the interest on this species and since then, abundance of *C. leptum* group, and particularly *F. prausnitzii*, has been extensively described in different intestinal disorders and metabolic diseases. Table 1.5 summarizes the studies performed by different methods, based in faecal or biopsy samples, which have reported changes in the abundance and/or prevalence of *F. prausnitzii* in different human diseases.

It is well established by molecular studies performed both in faecal or mucosa-associated communities, that in CD patients numbers of *F. prausnitzii* are depleted (Table 1.5). Several studies agree that this feature is characteristic of those patients with CD with ileal involvement, and presumably low numbers of this species may worsen disease progression. In this sense, Sokol *et al.* (2008) reported that I-CD patients with low *F. prausnitzii* abundance had a higher risk of post-operative recurrence, mostly having a relapse six months after operation [118]. However, concerning C-CD patients, the reduction in *F. prausnitzii* levels is still unclear, with some studies reporting that there is no reduction [87] or even a slightly increased prevalence [115]. A recent study based on paediatric CD patients, treatment naïve, indicates that this species abundance is higher in this subgroup of patients suggesting a more complex role for *F. prausnitzii* than initially thought [241]. Altogether, these findings indicate that in future studies it is important to take into account disease location of the patients, activity status, as well as medication in order to reach a consensus on *F. prausnitzii* role in CD.

Results in UC patients are controversial. Some works report a reduction in *F. prausnitzii* load [15, 112, 242-244], whereas others do not observe any difference in *F. prausnitzii* abundance in comparison to the control group [85, 206, 241] (Table 1.5). Sokol *et al.* (2009) reported that *F. prausnitzii* abundance is reduced only in active UC patients [117], which highlights the importance of considering this clinical data in future studies.

Table 1.5. Variation of *F. prausnitzii* prevalence and/or abundance in faecal or mucosal samples from subjects with different intestinal disorders in comparison to the respective control group in each study. Data using different methods has been collected. (Adapted from [225]).

Disorder	Samples	Technique	n	Mean ages	*F. prausnitzii* variation	Ref
CD	faeces	TGGE, qPCR	20	31.2(±14.1)	↓	[112]
CD	faeces	PCR	20	ND	↓	[206]
CD	faeces	T-RFLP	67	30.1 (±11.6)	↓	[150]
CD	faeces	PCR-DGGE, qPCR	68	45 (25-76)	↓	[83]
CD	faeces	Microarray, qPCR	16	31 (25–39)	↓	[86]
CD	faeces	qPCR	47	35.3(±9.4)	↓	[246]
CD	faeces	FISH	82	34.8(17-78)	↓	[114]
CD	faeces	FISH	50	39(19-68)	↓	[247]
CD	faeces	FISH	28	44.3(21-76)	↓	[248]
CD	faeces	FISH	50	39.0 (19-68)	↓	[247]
CD	biopsies	PCR-DGGE	19	36.7 (±3.72)	↓	[85]
CD	biopsies (colonic)	Pyrosequencing, qPCR	13	12.2(8.0-16.3)	↑	[241]
CD	biopsies	FISH	20	32.5	↓	[15]
A-CD	faeces	qPCR	22	36.9 (±3.3)	↓	[117]
A-CD	faeces	qPCR	21	ND	↓	[237]
A-CD	faeces	FISH	101	39 (14)	↓	[249]
A-CD	mucosa associated	FISH	ND	ND	↓	[118]
A-CD	mucosa associated	qPCR	ND	ND	↓	[250]
A-CD	faeces	qPCR	ND	ND	↓	[250]
R-CD	faeces	GoArray, qPCR	6	31(18-44)	↓	[113]
R-CD	faeces	qPCR	10	39.1 (±4.2)	→	[117]
R-CD	faeces	qPCR	19	ND	↓	[237]
I-CD	faeces and biopsies	PCR-DGGE	18	35.2 (18-58)	↓	[251]
I-CD	biopsies	cloning	7	45.3 (±18)	↓	[115]
I-CD	mucosa associated	qPCR	6	50.8 (±4.5)	↓	[87]
C-CD	biopsies	cloning	7	52.6 (±18)	↑	[115]
C-CD	mucosa associated	qPCR	8	49 (±18.5)	→	[87]
UC	faeces	*In vitro* M-SHIME	6	40.5 (33-78)	↓	[244]
UC	faeces	PCR	14	ND	→	[206]
UC	faeces	TGGE, qPCR	22	38.4 (±11.3)	↓	[112]
UC	faeces	qPCR	127	43 (32-55)	↓	[88]
UC	faeces	FISH	105	41.2 (18-84)	↑	[114]
UC	biopsies	PCR-DGGE	2	31 (±2.37)	→	[85]
UC	biopsies (colonic)	PCR-cloning 16S rRNA gene	8	51 (19-63)	↓	[243]
UC	biopsies (colonic)	Pyrosequencing, qPCR	12	13.0 (8.5-15.8)	→	[241]
UC	biopsies (colonic)	Sequencing 16S rRNA	1	12	↑	[143]
UC	biopsies	FISH	20	44.4	↓	[15]
A-UC	faeces	qPCR	13	39.7 (±3.5)	↓	[117]
A-UC	faeces	qPCR	13	ND (paediatric)	→	[237]
A-UC	mucosa associated	qPCR	ND	ND	↓	[250]
A-UC	faeces	qPCR	ND	ND	↓	[250]
R-UC	mucosa associated	qPCR	ND	ND	→	[250]
R-UC	faeces	qPCR	ND	ND	→	[250]
R-UC	faeces	qPCR	116	40 (32–46)	↓	[252]
R-UC	faeces	qPCR	4	35 (±4.3)	→	[117]
R-UC	faeces	qPCR	16	ND (paediatric)	→	[237]
Pouchitis	biopsies	PCR-cloning 16S rRNA gene	8	39 (19-64)	→	[243]
Pouchitis FAP	biopsies	PCR-cloning 16S rRNA gene	3	32 (30-54)	→	[243]
IdC	faeces	FISH	17	46.4 (19–80)	→	[114]
IC	biopsies	PCR-DGGE	1	27	→	[85]
SLC	biopsies	FISH	20	37.3	↓	[15]
SLC	faeces	FISH	9	32 (19–66)	→	[114]
IBS	faeces	PCR	21	ND	→	[206]
IBS	faeces	pyroseqencing,microarray, qPCR	62	49 (22-66)	↓	[176]
IBS	faeces	pyroseqencing,microarray, qPCR, FISH	22	13.2 (8-18)	→	[253]
IBS	faeces	cloning, qPCR	24	47.3 (21–65)	→	[169]
IBS	faeces	qPCR	13	39.7 ± 7	→	[254]
IBS	faeces	qPCR	27	46.5 (20-65)	→	[76]
IBS	faeces	FISH	45	45.4(24–72)	→	[114]
IBS	biopsies	FISH	20	47.8	→	[15]
Chronic diarrhoea	faeces	FISH	30	46.1 (21-77)	↓	[247]
CRC	faeces	Pyrosequencing, qPCR	60	67.1(±11.6)	→	[190]
CRC	faeces	qPCR	20	51.5 (18-77)	↓	[255]
NET	faeces	FISH	66	58.5 (27-85)	↓	[247]
UGC	faeces	qPCR	9	54 (40-67)	→	[187]
Celiac disease	faeces	FISH	12	32 (22–68)	↓	[114]
Celiac disease	faeces	FISH	24	5.5 (2.1-12.0)	↓	[245]
Obesity (Chinese)	faeces	qPCR	54	46.0 (42.0-50.0)	→	[256]
Obesity (French)	faeces	qPCR	23	42 (± 2)	↓	[242]
Obesity (Indian)	faeces	qPCR	15	13 (10-15)	↑	[255]
Obesity (Swiss)	faeces	qPCR	15	10.6 (8-14)	→	[257]
Obesity & type 2 diabetes	faeces	qPCR	7	49 (± 5)	↓	[242]
Obesity & type 2 diabetes	faeces	metagenomic analysis	6	46.1 (38-53)	↓	[258]
Appendicitis	Ressected tissue	FISH	70	ND	↓	[259]

H, healthy subjects; CD, Crohn's disease; A-CD, active CD; R-CD, remission CD; I-CD, ileal CD; C-CD, colonic CD; UC, ulcerative colitis; A-UC, active UC; R-UC, remission UC; FAP, familial adenomatous polyposis; IdC, indetermined colitis; IC, ischemic colitis; SLC, self-limiting colitis; IBS, irritable bowel syndrome; CRC, colorectal cancer; NET, neuroendocrine tumour of the midgut; UGC, upper gastrointestinal cancer; PCR-DGGE, Polymerase Chain Reaction and Denaturing Gradient Gel Electrophoresis; T-RFLP, Terminal Restriction Fragment Length Polymorphism; FISH, Fluorescent In Situ Hybridization; qPCR, quantitative Polymerase Chain Reaction; ND, not determined; ↑, increase of *F. prausnitzii* in comparison to the control group; ↓, decrease of *F. prausnitzii* in comparison to the control group; → no statistically significant differences in *F. prausnitzii* in comparison to the control group of the respective study.

Nevertheless, it seems likely that *F. prausnitzii* and its anti-inflammatory properties may also play a protective role in UC patients as it has been reported that low *F. prausnitzii* abundance in patients with ileal pouch have a higher risk to have pouchitis [243]. However, it still

remains to be elucidated if these patients already had lower *F. prausnitzii* abundance, and therefore underwent colectomy, or if it is due to the operation that *F. prausnitzii* abundance has diminished.

Another kind of chronic inflammatory disorder affecting exclusively the small intestine is celiac disease, in which genetically predisposed individuals feature a permanent intolerance to dietary gluten. *F. prausnitzii* abundance has been shown to be reduced in both untreated and treated celiac disease children compared with controls [79, 80]. A reduction on *F. prausnitzii* has also been observed when healthy adults are enrolled in a gluten-free diet [245]. These results suggest that polysaccharides intake may be playing a role, since these dietary compounds usually reach the distal part of the colon, and constitute one of the main energy sources for beneficial components of gut microbiota. However, a link between the activation of the adaptive and innate immune response as responsible for the reduction on *F. prausnitzii* abundance cannot be ruled out, as celiac children feature a lower relative abundance of this bacterium before any nutritional intervention [245].

No differences in *F. prausnitzii* abundance have been reported in other inflammatory disorders such as indeterminate colitis (IdC) [114], ischemic colitis (IC) [85] and self-limiting colitis (SLC) [15, 114, 247] whereas no conclusive results are available on irritable bowel syndrome (IBS) (Table 1.5). Initial studies conducted on IBS patients did not report any reduction on this species abundance either in biopsies or faecal samples [15, 114, 206, 254]. More recent studies based on deeper molecular analyses have nevertheless evidenced that *Faecalibacterium*-related bacteria are lower in IBS patients, particularly in those with IBS alternating type [176]. In contrast, no changes have been observed in diarrhoea-predominant IBS patients, nor in constipation-predominant IBS patients, what suggests that only a subtype of the disease is associated with reduced *F. prausnitzii* abundance. Taken together these findings indicate that the abundance of *F. prausnitzii* might be a reliable indicator of dysbiosis in IBD patients' although in some cases other intestinal disorders, that share some features with IBD such as micro-inflammation in IBS, can also have a diminished numbers of this species.

Concerning other intestinal disorders *F. prausnitzii* abundance has also been analyzed in colorectal cancer (CRC), neuroendocrine tumours (NET), upper gastrointestinal cancer (UGC) [187], chronic idiopathic diarrhoea and appendicitis. Although the reported changes are not always consistent, it seems that *F. prausnitzii* numbers are also lower in CRC patients [190, 255]. A remarkable *F. prausnitzii* depletion has been found in the stool of patients with NET of the midgut [247], whereas no changes have been observed in patients with NET of either the foregut or the hindgut. No reduction of *F. prausnitzii* abundance was found in UGC patients either [187]. Patients suffering of chronic idiopathic diarrhoea also have

reduced *F. prausnitzii* numbers in stools when compared to healthy controls [247]. Finally, *F. prausnitzii* abundance has been inversely related to the severity of appendicitis [259].

Finally, there are inconsistencies between different studies in elucidating *F. prausnitzii* role in metabolic disorders such as obesity. It has been reported that the Firmicutes to Bacteroidetes ratio is altered in overweight and obese subjects. On the one hand, the *Firmicutes* have been shown to be significantly higher in obese adult individuals [38, 62] and in Indian obese children [255]. The authors suggested that high numbers of *F. prausnitzii* leads to higher energy intake, because *F. prausnitzii* is responsible for a significant proportion of fermentation of unabsorbed carbohydrates in the gut [255]. In contrast, the presence of *F. prausnitzii* has also been linked to the reduction of low-grade inflammation in obesity and diabetes independently of caloric intake [242, 258].

1.3.7. Diet and prebiotics influence on *F. prausnitzii* population

It is not clear yet how *F. prausnitzii* population can be modulated in the gut and few studies have been performed in this sense. It can be hypothesized that diet and some medication can influence this species numbers in the colon although the exact mechanism or the crucial factors for its stimulation remain to be elucidated.

Concerning the link between the amount of *F. prausnitzii* in the healthy human gut and diet, Mueller and co-workers (2006) studied country-related differences in faecal microbiota and observed that the level of *F. prausnitzii* was highest in the Swedish study group compared to Italian, German and French cohorts, and suggested that this could be due to dietary habits [34]. It has been reported that *F. prausnitzii* abundance is significantly reduced during both fibre-free and fibre-supplemented diets provided through liquid enteral formulas [48]. Whereas the reduction of *F. prausnitzii* in the fibre-free diet was attributed to the lack of this compound, the no-stimulation of *F. prausnitzii* population in the fibre-supplemented diet could be explained because *F. prausnitzii* growth may not be supported by pea fibre, which was the only fibre source in this study. In line with these findings, the no stimulation of *F. prausnitzii* by a fibre-free exclusive enteral nutrition in CD patients has been recently reported [260]. Besides, no change on *F. prausnitzii* relative abundance was observed in healthy obese subjects enrolled in a carbohydrate-restricted diet [47].

On the contrary, inulin-derived prebiotics have been shown to significantly increase *F. prausnitzii* concentration in the gut [55, 261]. In addition, gluten and high energy intakes in obese people have been correlated with changes in *F. prausnitzii* numbers, suggesting that polysaccharides may have an effect in modulating *F. prausnitzii* population in the colon. This can be performed either in a direct manner (i.e. carbohydrate fermentation by *F. prausnitzii*) or indirectly, if they boost other species that interact with *F. prausnitzii* in the gut. A co-

ocurrence analysis based on a relative abundance matrix previously reported in Qin *et al.*, 2010 [26] has evidenced that *F. prausnitzii* interacts with several members of the Bacteroidetes and the *Clostridium* Cluster XIVa (*C. coccoides* group) [235].

1.3.8. Effects of medication on *F. prausnitzii* population

Concerning *F. prausnitzii* sensitivity to antibiotics few studies have been published to date. *F. prausnitzii* group abundance is significantly reduced in stools of patients receiving digestive microbial decontamination in intensive care unit [262]. Among the antibiotics delivered to patients during this treatment, it seems most likely that *F. prausnitzii* is suppressed by tobramycin since the minimum inhibitory concentration (MIC) was of 4μg/ml, a concentration easily reached with the intestinal decontamination regime. In contrast, sensitivity tests showed that this species can resist higher concentrations of cefotaxime (MIC >32μg/ml) and colistin (MIC >256 μg/ml). The use of rifaximin, reported to induce clinical remission in patients with active CD, has been associated with an increased level of *F. prausnitzii* [263], although this study does not allow elucidating if this is due to restoration in gut homeostasis of CD patients or if it is due to the suppression of species that compete with *F. prausnitzii*. Recently, a study on *Faecalibacterium* sp. isolates from calves and piglets has evidenced that most of them are resistant to ciprofloxacin and sulfamethoxazole-trimethoprim. Besides, over a 50% of the isolates showed resistance to tetracycline, amikacin, cefepime and cefoxitin, and some of them also featured multidrug resistance [264].

In addition, some other studies have reported data concerning medication effect on *F. prausnitzii* population. To date, some IBD treatments have been already reported to have a positive impact on *F. prausnitzii* population. For instance, Infliximab or a high dose cortisol therapy can completely restore *F. prausnitzii* concentrations from non detectable numbers to levels higher than 1.4×10^{10} bacteria/ml within days [114]. Other specific treatments such as chemotherapy and interferon α-2b were shown to reverse the depletion of *F. prausnitzii* in patients with midgut NET, whereas somatostatin analogs had no influence on this species [247]. These results suggest that restoration of the gut conditions thanks to medication can have an effect on counterbalancing *F. prausnitzii* depletion in a diseased intestine.

2 scope of the thesis

Despite being one of the three most abundant species found in the human gut, and an active member of the microbiome, little attention was paid to *F. prausnitzii*, a member of the phylum Firmicutes (Ruminococcaceae) until very recently. The rising number of studies reporting *F. prausnitzii* depletion in patients suffering intestinal disorders, and pointing out its beneficial role to maintain gut health has prompted interest in this species in the last few years. However, information about its growth requirements, the genetic diversity comprised within this species, and how its abundance is affected by intestinal disorders remains unknown. These have been the three main topics addressed in this Thesis.

Because information on genetic diversity and substrate utilization of *F. prausnitzii* is limited, mainly due to its difficulty to be cultured, in **Chapter 1** a phylogenetic and phenotypic characterisation of *F. prausnitzii* isolates from human faeces has been performed. The phylogenetic characterisation of *F. prausnitzii* is of interest in order to know the genetic diversity within this species and to which extent the cultured isolates are representative of the diversity found in *in vivo*. Besides, the phenotypic characterisation of *F. prausnitzii* strains isolated from subjects of different age, ethnic group and under different diets is of interest in order to gain knowledge of the intraspecies phenotypic diversity, and to determine which are the nutritional requirements and crucial factors for its survival in the gut. Therefore, in this first part of the Thesis it was examined the phylogeny, substrate utilisation and influence of

gut environmental factors on growth of recent *F. prausnitzii* strains isolated from healthy subjects.

Phylogenetic analysis in the first part of the study showed that mainly two *F. prausnitzii* phylogroups, which include the current cultured representatives, were responsible for the abundance of this species in the gut of healthy subjects. In addition, all cultured *F. prausnitzii* representatives were extremely sensitive to small changes in the physico-chemical conditions of the colonic environment that may occur in gut disease, but some differences between isolates were observed. Because information about the genetic diversity within this species is missing, it was wondered if *F. prausnitzii* populations hosted by healthy subjects are different from that found in patients suffering gut disorders. Therefore, **Chapter 2** of this Thesis was addressed to study the microdiversity of *F. prausnitzii* in healthy and diseased gut.

Finally, **Chapter 3** was addressed to determine the abundance of mucosa-associated *F. prausnitzii* and to assess its usefulness as diagnostic biomarker in inflammatory bowel diseases. By the time to start this Thesis, mounting evidences suggested that *F. prausnitzii* abundance is depleted in Crohn's disease patients and in the following years a depletion of this species load was reported in other intestinal disorders. Besides, other species such as *Escherichia coli* were consistently reported as dysbiosis-signature of Crohn's disease patients. The results in the second Chapter of this Thesis evidenced that the imbalance in the phylogroups of *F. prausnitzii* could be a signature of gut disease. The use of gut microbiota species as biomarkers to diagnose or prognose intestinal diseases is an interesting approach that may fill two major unmet clinical needs. On the one hand, they may be a useful source of additional information for disease diagnosis. Besides, the application of gut microbiota as prognostic tool is of interest in chronic relapsing diseases such as inflammatory bowel disease to monitor patients over time and guide their management. However, comprehensive studies taking into account patients clinical features such as disease location, activity status as well as medication, are scarce. In addition, the abundance of *F. prausnitzii* phylogroups in gut health and disease is unknown. This third part of this Thesis aimed at determining if *F. prausnitzii* can be of assistance in inflammatory bowel disease diagnostics and prognostics.

Altogether, this work offers a journey from classical microbiology to applied microbiology which has allowed gaining insight into this species ecophysiology and phylogeny in healthy and diseased gut, and to elucidate its applicability as biomarker to assist in inflammatory bowel disease diagnosis.

3 aims

The main aim of this Thesis is to gaining insight into *F. prausnitzii* physiology, diversity and abundance in healthy and diseased gut, with the further aim to apply it in intestinal disorders diagnosis and/or prognosis. To address these challenges, both culture dependent and molecular techniques have been used, and the results are presented in the following section organised into three Chapters (which comprise four papers), with the following specific objectives:

1. To examine the phylogeny, phenotypic characteristics, and influence of gut environmental factors on growth of *F. prausnitzii* strains isolated from healthy subjects (Article I).

2. To determine if subjects with gastrointestinal disease host different mucosa-associated *F. prausnitzii* populations from healthy (Article II).

3. To quantify mucosa associated *F. prausnitzii* and to establish its potential usefulness as biomarker in gut diseases diagnostic and/or prognostic (Article III and Article IV).

The results of this PhD Thesis have been published in scientific journals included in the Journal Citation Report of the Institute of Scientific Information. Therefore, this PhD thesis has been presented as a compendium of publications. According to the current rules that regulate this format type at the Universitat de Girona, for the PhD programme Experimental Sciences and Sustainability a minimum of two articles published or accepted are required. At least one article from the first quartile and the other from the second quartile of the Journal Citation Report of the Institute of Scientific Information, both of which the doctoral candidate must be the first author. However all the articles obtained during the PhD period of the candidate have been included in this work for coherence of the research project performed.

Chapter 1‡

Phylogenetic characterisation, substrate utilization and tolerance to gut environmental factors of *Faecalibacterium prausnitzii* isolates from human faeces

Significance of this study

What was already known on this subject?

- *Faecalibacterium* (formerly *Fusobacterium*) *prausnitzii* is one of the three most abundant species detected in human faeces but following its first isolation, this species received little attention, partly because of its oxygen sensitivity.
- *F. prausnitzii* is a non-motile, non-flagellated bacterium that produces butyrate, D-lactate and formate, and utilizes acetate during glucose fermentation. Phenotypic characterisation of four isolates revealed that it can grow on fructose, fructo-oligosaccharide, starch and inulin; whereas none can utilize arabinose, melibiose, raffinose, rhamnose, ribose and xylose. The isolates differed in their ability to ferment cellobiose, maltose and melezitose.
- The relative abundance of *F. prausnitzii* among the human colonic microbiota, as estimated by 16S rRNA-based culture-independent methods, is reduced in certain forms of inflammatory bowel disease.

What are the new findings?

- The cultured strains are representative of *F. prausnitzii* sequences detected by direct analysis of faecal DNA from healthy subjects, and separated the available isolates into two phylogroups.
- *F. prausnitzii* is a metabolically versatile bacterium capable to ferment carbohydrates of different structure and origin found in the gut.
- All strains tested were bile sensitive, while inhibition at mildly acidic pH was strain dependent.

How might this impact on research in the foreseeable future?

- These attributes help to explain the abundance of *F. prausnitzii* in the colonic community but also suggest factors in the gut environment that may limit its distribution.

---

‡

This chapter has been published in the journal *Applied and Environmental Microbiology*: Lopez-Siles, M *et al.* **Cultured representatives of two major phylogroups of human colonic *Faecalibacterium prausnitzii* can utilize pectin, uronic acids and host-derived substrates for growth.** *Appl. Environ. Microbiol.* 2012; 78 (2): 420-428.

Article I

AEM Journals.ASM.org

Cultured Representatives of Two Major Phylogroups of Human Colonic *Faecalibacterium prausnitzii* Can Utilize Pectin, Uronic Acids, and Host-Derived Substrates for Growth

Mireia Lopez-Siles,[a] Tanweer M. Khan,[b] Sylvia H. Duncan,[c] Hermie J. M. Harmsen,[b] L. Jesús Garcia-Gil,[a] and Harry J. Flint[c]

Grup de Microbiologia Clínica i Malalties Infeccioses, Departament de Biologia, Facultat de Ciències, Universitat de Girona, Girona, Spain[a]; Department of Medical Microbiology, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands[b]; and Microbial Ecology Group, Rowett Institute of Nutrition and Health, University of Aberdeen, Aberdeen, United Kingdom[c]

Faecalibacterium prausnitzii* is one of the most abundant commensal bacteria in the healthy human large intestine, but information on genetic diversity and substrate utilization is limited. Here, we examine the phylogeny, phenotypic characteristics, and influence of gut environmental factors on growth of *F. prausnitzii* strains isolated from healthy subjects. Phylogenetic analysis based on the 16S rRNA sequences indicated that the cultured strains were representative of *F. prausnitzii* sequences detected by direct analysis of fecal DNA and separated the available isolates into two phylogroups. Most *F. prausnitzii* strains tested grew well under anaerobic conditions on apple pectin. Furthermore, *F. prausnitzii* strains competed successfully in coculture with two other abundant pectin-utilizing species, *Bacteroides thetaiotaomicron* and *Eubacterium eligens*, with apple pectin as substrate, suggesting that this species makes a contribution to pectin fermentation in the colon. Many *F. prausnitzii* isolates were able to utilize uronic acids for growth, an ability previously thought to be confined to *Bacteroides* spp. among human colonic anaerobes. Most strains grew on *N*-acetylglucosamine, demonstrating an ability to utilize host-derived substrates. All strains tested were bile sensitive, showing at least 80% growth inhibition in the presence of 0.5 μg/ml bile salts, while inhibition at mildly acidic pH was strain dependent. These attributes help to explain the abundance of *F. prausnitzii* in the colonic community but also suggest factors in the gut environment that may limit its distribution.

Faecalibacterium (formerly *Fusobacterium*) *prausnitzii* (11) is one of the three most abundant species detected in human feces by anaerobic cultivation (32) and by 16S rRNA-based molecular analyses (21, 51, 52, 57). Following its first isolation (4, 20), this species received little attention, partly because of its oxygen sensitivity (14), until new isolates became available from studies on the dominant butyrate-producing bacteria from the human colon (2) that allowed the definition of the new genus *Faecalibacterium* (11). Interest in this bacterium has increased recently with reports that the relative abundance of *F. prausnitzii* among the human colonic microbiota, as estimated by 16S rRNA-based culture-independent methods, is reduced in certain forms of inflammatory bowel disease. Crohn's disease (CD) patients, mainly those with ileal involvement, have been reported to exhibit diminished prevalence of *Firmicutes*, often with a concomitant increase in *Proteobacteria* (15, 30, 60). Molecular analysis of both fecal and biopsy samples has revealed that the depletion in the former is due in part to decreased abundance of the *F. prausnitzii* group (6, 45, 47, 50, 60). Reduced *F. prausnitzii* abundance has also been reported in colorectal cancer (1) and in the frail elderly (29, 56), leading to the suggestion that this bacterium could provide an indicator of a healthy gut microbiota. *F. prausnitzii* is one of the main sources of butyrate in the colon (27, 37), and the multiple effects of butyrate as the preferred energy source for the colonocytes and upon apoptosis, inflammation, and oxidative stress are generally considered to be beneficial to intestinal health (18, 37, 40). *F. prausnitzii* is also thought to have additional anti-inflammatory properties that are suggested by cellular studies and trinitrobenzene sulfonic acid colitis models in mice (49).

In view of the proposed role of *F. prausnitzii* in intestinal health, it is important to gain a better understanding of the microbial ecology of this species. It is currently unclear what major substrates, of dietary or host origin, are likely to support growth and what factors in the gut environment may influence its distribution in the intestine. It is also important to establish how much genetic and phenotypic variation occurs within this species and the extent to which available cultured strains represent the diversity present *in vivo*. This study addresses these questions by examining the characteristics of the available cultured strains, including new isolates from healthy humans.

MATERIALS AND METHODS

Bacterial strains and growth conditions. The *F. prausnitzii* isolates listed in Table 1 were from stocks held by the authors (S. H. Duncan, Rowett Institute of Nutrition and Health, Aberdeen, United Kingdom, and H. J. M. Harmsen, Department of Medical Microbiology, University of Groningen, Groningen, The Netherlands), and all are of human fecal origin (Table 1). *F. prausnitzii*-related isolates were obtained from the highest countable dilution of human fecal samples in roll tubes of anaerobic M2GSC medium (31), as described previously (2). Anaerobic culture methods were those of Bryant (3) using Hungate culture tubes, sealed with butyl rubber septa (Bellco Glass). Additional *F. prausnitzii* strains designated HTF isolates were isolated from freshly voided human stools,

Received 13 September 2011 **Accepted** 9 November 2011
Published ahead of print 18 November 2011
Address correspondence to Harry J. Flint, H.Flint@abdn.ac.uk.
M. Lopez-Siles and T. M. Khan contributed equally.
Supplemental material for this article may be found at http://aem.asm.org/.

doi:10.1128/AEM.06858-11

TABLE 1 Details of *F. prausnitzii* strains included in this study[a]

Isolate code	Laboratory of isolation	Volunteer	Sex	Age (yr)	Culture collection	Reference(s) for original isolation
ATCC 27768		1	Unknown	Unknown	ATCC 27768	4
A2-165	RINH	2	F	34	DSMZ 17677	2, 11
L2-15	RINH	3	M	2		2
L2-39	RINH	3	M	2		2
L2-6	RINH	3	M	2		2, 11
L2-61	RINH	3	M	2		2
M21/2	RINH	4	F	36		26
S3L/3	RINH	5	F	46		26
S4L/4	RINH	5	F	46		26
HTF-A	GU	6	M	31		This study
HTF-B	GU	6	M	31		This study
HTF-C	GU	6	M	31		This study
HTF-E	GU	7	M	44		This study
HTF-F	GU	7	M	44		This study
HTF-I	GU	8	M	28		This study
HTF-60C	GU	8	M	28		This study
HTF-75H	GU	9	M	65		This study

[a] All the isolates were obtained from human fecal samples of healthy volunteers consuming omnivorous diets. Abbreviations: RINH, Rowett Institute of Nutrition and Health, Aberdeen (Scotland), United Kingdom; GU, Groningen University, Groningen, The Netherlands; F, female; M, male.

by plating 1 μl of the fecal material with a loop as a lawn directly on YCFAG medium (see below). After 12 h to 16 h of incubation at 37°C in an anaerobic tent (80% N_2, 12% CO_2, and 8% H_2), 500 translucent colonies per sample were selected and subcultured on fresh plates (50 per plate in a grid-like fashion). After growth, the colonies were presumptively identified based on morphology, eliminating 95% of the colonies. The remaining colonies were further purified and Gram stained. Up to 5 colonies per sample were finally identified by 16S rRNA gene sequencing. The isolates were routinely maintained by being grown for 16 to 18 h at 37°C in 7.5-ml aliquots of M2GSC medium (31) and maintained anaerobically using O_2-free CO_2. The low-percent G+C Gram-positive *Firmicutes* strains screened for pectin utilization in this study (see Table 3) were also from stocks held by the authors (Rowett Institute of Nutrition and Health, Aberdeen, United Kingdom), and several came from previous studies (2, 26). The strains included *Roseburia intestinalis* L1-82 (DSM 14610^T), *Roseburia hominis* A2-183 (DSM 16839^T), *Roseburia inulinivorans* strains A2-194 (DSM 16841^T) and L1-83, *Roseburia faecis* M72/1 (DSM 16841^T) and M88/1, and *Eubacterium rectale* A1-86 (DSM 17629), M104/1, and L2-21, with type strains deposited with the Deutsche Sammlung von Mikrooganismen und Zellkulturen (DSMZ). Other *Firmicutes* tested in the study included *Butyrivibrio fibrisolvens* 16/4, which was isolated as a butyrate-producing wheat bran degrader (41). *Eubacterium siraeum* 70/3 (8) and V10Sc8a are also isolates from human fecal samples. *Eubacterium eligens* DSM 3376 was from DSMZ, *Bacteroides thetaiotaomicron* B5482 was a gift from A. Salyers, and both strains were included in the coculture studies.

Growth medium. YCFA medium consists of (per 100 ml) Casitone (1.0 g), yeast extract (0.25 g), $NaHCO_3$ (0.4 g), cysteine (0.1 g), K_2HPO_4 (0.045 g), KH_2PO_4 (0.045 g), NaCl (0.09 g), $(NH_4)_2SO_4$ (0.09 g), $MgSO_4 \cdot 7H_2O$ (0.009 g), $CaCl_2$ (0.009 g), resazurin (0.1 mg), hemin (1 mg), biotin (1 μg), cobalamin (1 μg), *p*-aminobenzoic acid (3 μg), folic acid (5 μg), and pyridoxamine (15 μg). In addition, the following short-chain fatty acids (SCFA) are included (final concentrations): acetate (33 mM); propionate (9 mM); isobutyrate, isovalerate, and valerate (1 mM each). Cysteine is added to the medium following boiling and dispensed into Hungate tubes while the tubes are flushed with CO_2. After autoclaving, filter-sterilized solutions of thiamine and riboflavin are added to give final concentrations of 0.05 μg ml^{-1} of each. For some experiments, the Casitone content was decreased to 0.2%; this modified medium is referred to as YcFA. Carbohydrate or other energy sources were added as indicated, and the final pH of the medium was adjusted to 6.8 $\pm$ 0.1.

DNA extraction, PCR amplification, and DGGE fingerprinting. DNA was extracted and purified from 18-h-old cultures of *F. prausnitzii* strains grown on M2GSC medium by using the Wizard genomic purification kit (Promega Corporation, Madison, WI). 16S rRNA sequences were amplified using universal bacterial primers GC-357F (33) and 907R (34) to give an approximately 580-bp product flanking variable regions V3 to V5. PCR and denaturing gradient gel electrophoresis (DGGE) were carried out as previously reported (30).

16S rRNA gene amplification and sequencing. 16S rRNA genes were amplified using the universal bacterial primers 7F and 1510R (23) as described previously (12). PCR products were cleaned with the Wizard PCR product purification kit (Promega, Southampton, United Kingdom) and used to obtain bidirectional partial 16S rRNA gene sequences by using primers 7F, 519F, 519R, 916F, 916R, and 1510R (16, 23) on a Beckman capillary sequencer. All primers were obtained from Eurofins MWG.

16S rRNA gene sequence full-length construction and phylogenetic analysis. Sequences from cultured isolates were manually inspected in order to assess quality. Sequence editing and assembling were carried out using the BioEdit sequence alignment editor, version 7.0.9.0 (17). Sequences were then aligned in Mothur (http://www.mothur.org) (46) using the SILVA bacterial database as a reference alignment, available at the same source. Alignment was then imported into the ARB software package (28) loaded with the SILVA 16S rRNA-ARB-compatible database (SSURef-100, August 2009, available through the SILVA rRNA database project at http://www.arb-silva.de/) (36). For the detection of chimeric sequences, each sequence was checked manually in the alignment, and phylogenetic trees were screened for sequences with unrealistically long branches or unique branching sites. Cultured representatives from the *Ruminococcaceae* were included as reference, and *Eubacterium desmolans* was used to root the tree. Phylogenetic analyses of the 16S rRNA gene sequences were conducted using the ARB software package, using the neighbor-joining (NJ) method (42) and the Jukes-Cantor (JC) algorithm for distance analysis. Tree topologies were evaluated using maximum parsimony and maximum likelihood methods. No filters or masks were used when constructing the trees. Bootstrapping analysis (1,000 replicates) was done to test the robustness of the NJ-JC tree using PHYLIP (13).

To assess which *F. prausnitzii* phylogroups were represented by the isolates, representative sequences of 16S rRNA genes directly amplified from fecal DNA were included (boldfaced in Fig. 2). These uncultured sequences were aligned and processed as described above and then added

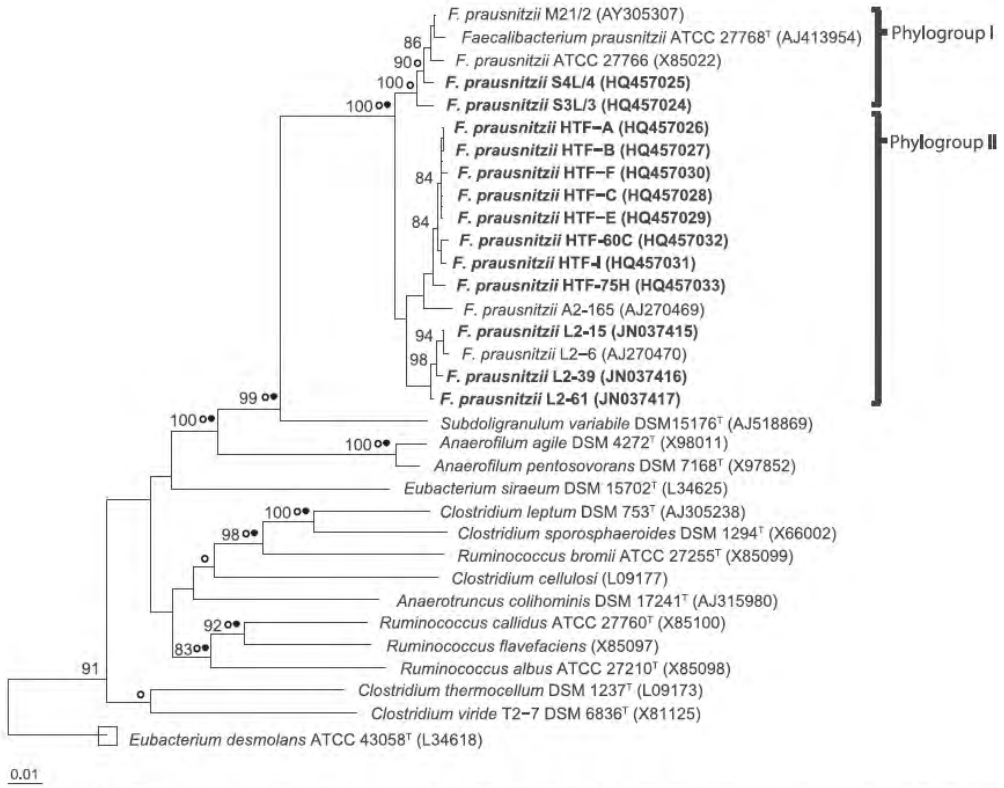


FIG 1 Phylogenetic relationship of *F. prausnitzii* isolates to other members of *Clostridium* cluster IV (*Ruminococcaceae*) based on 16S rRNA gene sequences. The tree was constructed using the ARB software package using the neighbor-joining method for distance analysis (Jukes-Cantor algorithm) with 1,533 informative positions considered (61 to 1,442 by *E. coli* 16S rRNA gene numbering). Bootstrap values above 80% (expressed as a percentage of 1,000 replications) are shown at branching points. Solid circles indicate branches that were consistent with calculations obtained by maximum-parsimony method. Empty circles represent those branches consistent with the maximum likelihood. The scale bar indicates the number of substitutions per site. *F. prausnitzii* isolates incorporated in this study are highlighted in bold. Sequence accession numbers are shown in parentheses. The database sequence for ATCC 27766 was included, but this strain was not studied here and it is not listed in Table 1.

to the isolate-based tree using the Parsimony Quick Add Marked Tool already implemented in the ARB software package, thereby maintaining the overall tree topology.

RAPD-PCR. Isolates were screened by random amplified polymorphic DNA PCR (RAPD-PCR) using the primer 1254, according to a previously described method (59). RAPD-PCR profiles were compared using the GelComparII software (Applied Maths, Belgium). The UPGMA (unweighted-pair group method using average linkages) method was used to build the dendrogram (see Fig. S1 in the supplemental material), and clusters were defined at a similarity score of >93.5%.

Carbohydrate utilization and assessment of bacterial growth. Substrate utilization was determined by adding a final concentration of 0.5% (wt/vol) sugar to YCFA medium. Where possible, growth was measured spectrophotometrically as optical density at 650 nm (OD_{650}) for triplicate cultures at regular intervals up to stationary phase. For insoluble xylan, however, fermentation was monitored by final pH measurement. To study competition for pectin, *F. prausnitzii* strains S3L/3 and A2-165 were inoculated individually and together with the known pectin-utilizing species *B. thetaiotaomicron* B5482 and *E. eligens* 3376 in cocultures and tricultures (see Table S2 in the supplemental material). These experiments used YcFA medium supplemented with 0.5% apple pectin (BDH Chemicals) that had been preadjusted to three different initial pH values (6.12, 6.45, and 6.79). Samples were collected at 0 h and 24 h to estimate bacterial numbers by fluorescent *in situ* hybridization (FISH), total sugar analysis, and SCFA concentrations. SCFA were analyzed by gas chromatography following conversion to *t*-butyldimethylsilyl derivatives (39). Total sugars were determined using the colorimetric phenol sulfuric assay (10).

Influence of initial pH and bile salts on bacterial growth. Each strain was inoculated into YCFA medium supplemented with 10 mM glucose (YCFAG) that had been adjusted to the three different initial pH values (6.7, 6.2, and 5.75) as described previously (12). Growth was followed for 24 h by measuring absorbance at 650 nm for triplicate cultures, and specific growth rates (h^{-1}) were calculated in exponential phase. The influence of bile salts (Sigma B8631) was assessed by inoculating culture into YCFAG medium containing 0% (control), 0.1%, 0.25%, or 0.5% bile salts (all percentages in wt/vol), in triplicate. Growth was measured spectrophotometrically up to 24 h using absorbance at the 650-nm wavelength. The pH of the medium was also monitored at the beginning and at the end of each experiment.

Enumeration of *F. prausnitzii* bacteria by FISH analysis. Cultures were prepared for analysis as described previously (19). Cell suspensions were applied to gelatin-coated slides. Dried slides were hybridized with 10 μl of the Fprau645 oligonucleotide probe (52) (50-ng/μl stock solution) and washed. Between 25 and 30 fields were counted per well using an epifluorescence microscope (Olympus) and image analysis software (Olympus Cell F digital imaging software) or manual counting for numbers of less than 10 fluorescent cells per field.

Statistical analysis. Quantitative parameters, such as growth rates and relative OD_{650}, were compared by one-way analysis of variance (ANOVA). The Bonferroni *post hoc* test was applied for multicomparisons of those variables with more than two subgroups of samples. Previously, data normality was assessed by the Shapiro-Wilks test and the Leven test was conducted to assess for homoscedasticity. The Kruskal-Wallis nonparametric test was performed when required. All statistical analyses were conducted via SPSS 15.0 (SPSS Inc., Chicago, IL).

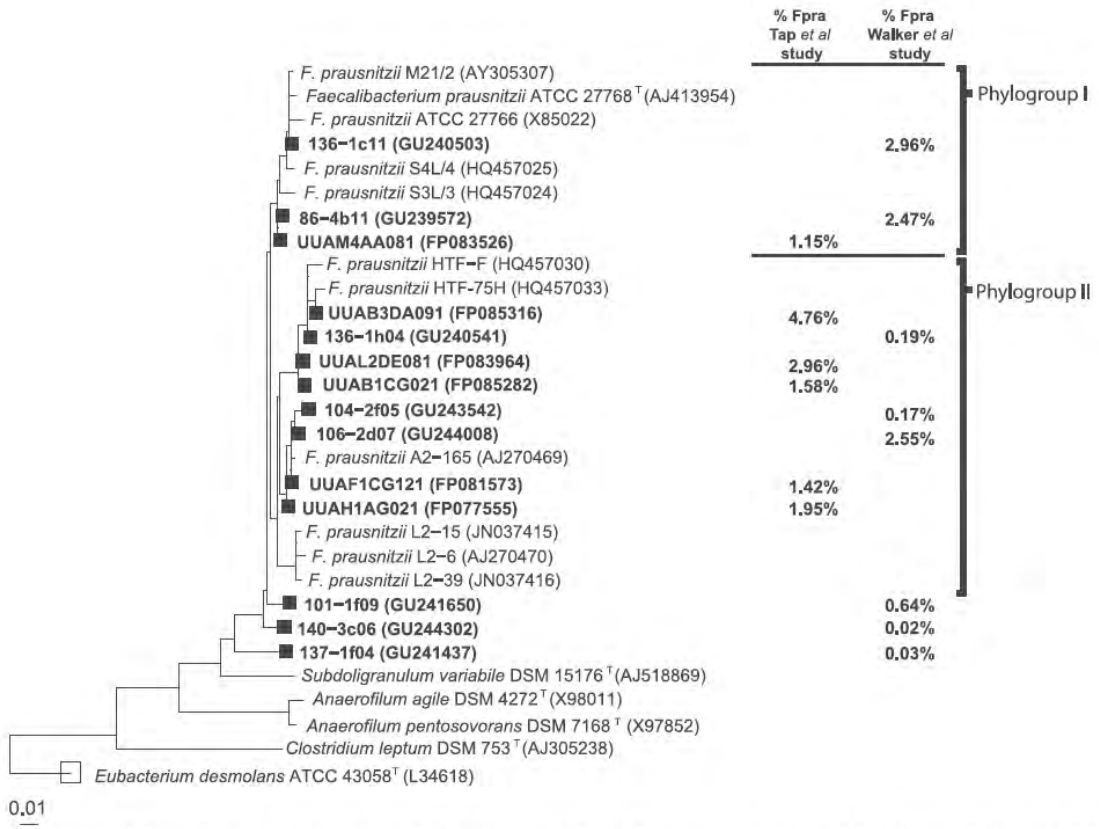


FIG 2 Neighbor-joining phylogenetic tree showing the relationship between cultured *F. prausnitzii* strains and directly amplified partial 16S rRNA gene sequences from human fecal samples. 16S rRNA sequence accession numbers are given in parentheses. Squares indicate OTU representative sequences from two recent studies on gut microbiota of healthy subjects (shown in boldface): the Tap et al. (53) study (1,443 *F. prausnitzii* sequences out of 10,456 clones from 17 healthy adults of both sexes) and the Walker et al. (57) study (534 *F. prausnitzii* sequences out of 5,915 total sequences from six obese males). The percentage of all clones represented by each OTU in each of these studies is shown on the right.

Nucleotide sequence accession numbers. The 16S rRNA gene full-length sequences of isolates S3L/3, S4L/4, HTF-A, HTF-B, HTF-C, HTF-E, HTF-F, HTF-I, HTF-60C, HTF-75H, L2-15, L2-39, and L2-61 were deposited in the GenBank/EMBL/DDBJ database under the accession numbers HQ457025 to HQ457033 and JN037415 to JN037417, respectively.

RESULTS AND DISCUSSION

Phylogenetic diversity of *Faecalibacterium prausnitzii*. Nearly full-length 16S rRNA gene sequences were determined for the first time here for 13 recent isolates of *Faecalibacterium prausnitzii* (Table 1; Fig. 1). The 16S rRNA sequences define two branches within the *Ruminococcaceae*, within which sequences share >97% sequence identity; these also include five sequences reported previously for the isolates M21/2, ATCC 27766, and ATCC 27768T (phylogroup I) and A2-165 and L2-6 (phylogroup II). The 18 isolates shown in Fig. 1 originated from 10 healthy individuals. Each of these 16S rRNA sequences is unique and came from a different colony, although there was a tendency for sequences to group by isolation and individual. This was also suggested by RAPD-PCR profiles for these strains (see Fig. S1 in the supplemental material). Comparison was also made with *F. prausnitzii*-related operational taxonomic units (OTUs) defined by partial 16S rRNA gene sequences obtained in two recent human studies by direct amplification from fecal DNA (53, 57) (Fig. 2). These

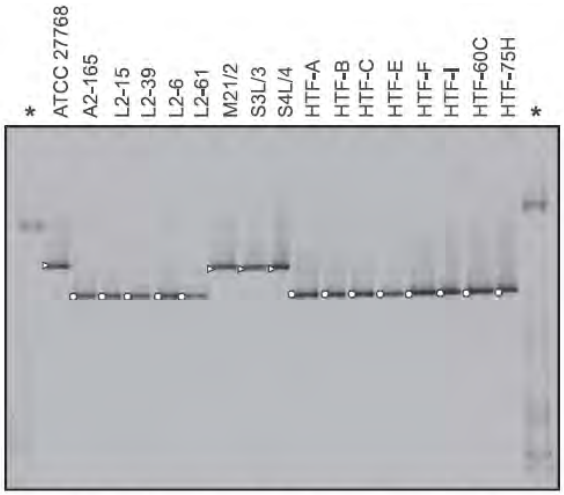


FIG 3 PCR-DGGE fingerprints from *F. prausnitzii* isolates. Isolates are distributed in two separate bands that correlate with phylogroup designation (△, phylogroup I; ○, phylogroup II). Asterisks indicate the ladder lanes (made by 16S rRNA gene fragments of *Mucor* sp., *Pseudomonas fluorescens*, and *Micrococcus luteus*, respectively, from the top to the bottom).

TABLE 2 Growth of *F. prausnitzii* strains on a range of carbohydrate substrates

Substrate[a]	Supplier and catalog no.	OD_{650} (mean ± SD) after 24 h									
		Phylogroup I strains				Phylogroup II strains					
		ATCC 27768	M21/2	S3L/3	S4L/4	A2-165	L2-15	L2-39	L2-6	HTF-75H	HTF-F
Glucose	BDH 10117	0.29 ± 0.02	0.96 ± 0.02	0.92 ± 0.18	0.83 ± 0.43	0.53 ± 0.13	0.29 ± 0.01	0.26 ± 0.05	0.32 ± 0.21	0.32 ± 0.02	0.85 ± 0.07
Cellobiose	Sigma C7252	0.26 ± 0.02	0.87 ± 0.33	0.81 ± 0.11	0.72 ± 0.21	0.63 ± 0.10	0.28 ± 0.01	0.18 ± 0.01	0.02 ± 0.07	0.32 ± 0.05	0.87 ± 0.01
Maltose	Sigma M5885	0.32 ± 0.35	0.85 ± 0.15	0.75 ± 0.07	0.82 ± 0.12	0.62 ± 0.07	0.44 ± 0.11	0.78 ± 0.05	0.22 ± 0.21	0.55 ± 0.10	1.01 ± 0.04
Rhamnose	Sigma R3875	—[b]	—	—	0.12 ± 0.03	—	—	—	—	—	—
Galacturonic acid	BDH 571670	0.12 ± 0.00	0.31 ± 0.04	0.45 ± 0.04	0.61 ± 0.06	0.21 ± 0.04	0.12 ± 0.01	0.07 ± 0.02	—	—	0.26 ± 0.02
Galactose	BDH G0750	0.24 ± 0.05	0.95 ± 0.03	0.44 ± 0.02	0.11 ± 0.09	0.80 ± 0.08	0.75 ± 0.28	—	0.61 ± 0.12	0.33 ± 0.28	0.66 ± 0.25
Pectin, apple	BDH 38052	0.31 ± 0.09	0.40 ± 0.04	0.36 ± 0.03	0.56 ± 0.02	0.66 ± 0.01	0.08 ± 0.00	0.07 ± 0.00	0.24 ± 0.02	0.18 ± 0.07	0.39 ± 0.07
Starch, potato	BDH 102713	—	0.06 ± 0.01	0.09 ± 0.02	—	0.07 ± 0.03	—	0.06 ± 0.01	0.07 ± 0.02	0.08 ± 0.06	0.05 ± 0.02
Inulin, chicory	Sigma I2255	0.21 ± 0.27	0.10 ± 0.00	0.08 ± 0.01	0.07 ± 0.01	0.80 ± 0.05	—	—	0.09 ± 0.18	0.18 ± 0.07	0.97 ± 0.26
Glucuronic acid	Fluka 71560	0.09 ± 0.00	—	0.28 ± 0.05	0.08 ± 0.01	0.83 ± 0.02	—	—	—	0.08 ± 0.03	0.17 ± 0.03
N-Acetylglucosamine	Sigma A8625	0.34 ± 0.03	0.88 ± 0.04	0.67 ± 0.00	0.57 ± 0.03	0.98 ± 0.01	0.18 ± 0.06	0.07 ± 0.00	—	0.20 ± 0.02	0.51 ± 0.24
Glucosamine HCl	BDH 962240	0.15 ± 0.01	0.31 ± 0.01	0.58 ± 0.01	0.34 ± 0.11	0.95 ± 0.15	0.13 ± 0.03	0.08 ± 0.02	0.14 ± 0.03	0.14 ± 0.03	0.16 ± 0.07

[a] None of the strains grew on arabinose, fucose, xylose, arabinogalactan, polygalacturonic acid, pectin (citrus), mucin (pig gastric), chondroitin sulfate, hyaluronic acid, and heparin. No growth was detected on xylan by final pH change. All substrates were obtained from Sigma.
[b] —, $\Delta OD_{650} < 0.05$. All values in the table were corrected for growth on basal medium without carbohydrate addition.

represent an additional 23 individuals. Phylogroups I and II together account for 97.9% of these directly amplified *F. prausnitzii*-related sequences, with phylogroup I more abundant in the six subjects examined by Walker et al. (57) (62%) than in the 17 subjects examined by Tap et al. (53) (8.3%).

DGGE analysis of PCR products amplified from phylogroup I isolates showed a distinct band position compared with phylogroup II isolates (Fig. 3). These band positions correspond to two dominant bands that have previously been associated with *F. prausnitzii* in DGGE analyses of 16S rRNA sequences amplified from human fecal and biopsy samples (22, 30). This previous work also suggested that there is a differential reduction in phylotypes related to M21/2 (phylogroup I) compared with A2-165 relatives (phylogroup II) in biopsy specimens (30) and fecal samples (22) from CD patients.

Substrate utilization by *Faecalibacterium prausnitzii* isolates. Growth on carbohydrates of dietary and host origin by four phylogroup I and six phylogroup II isolates is shown in Table 2.

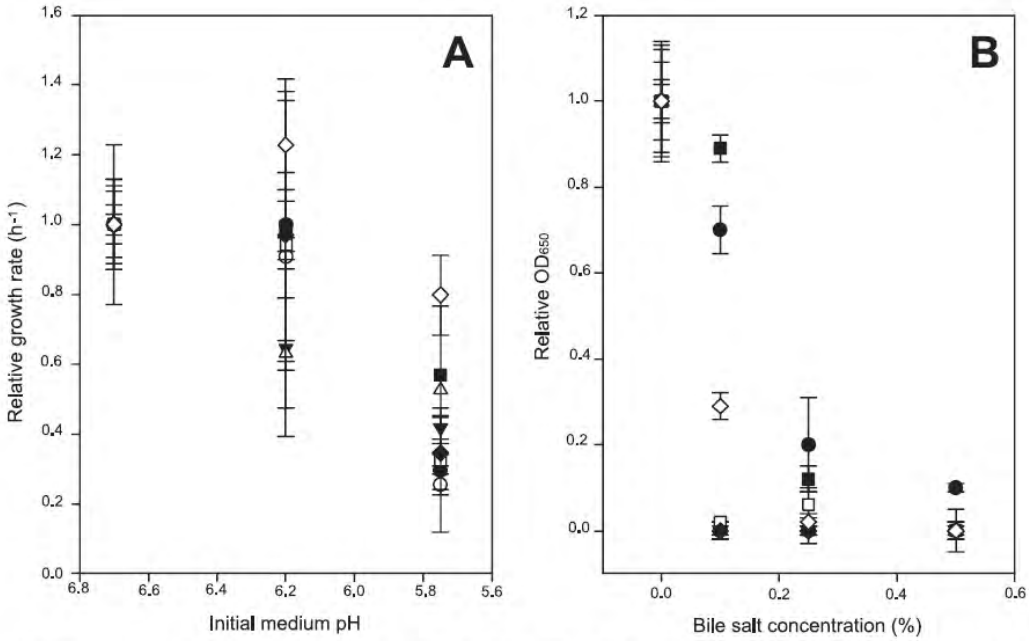


FIG 4 Tolerance of *F. prausnitzii* isolates to changes in initial medium pH values and bile salt concentrations. (A) Relative growth rates (h^{-1}) of *F. prausnitzii* strains on YCFAG medium at three initial pH values (6.7, 6.2, and 5.75) have been represented. For comparison, the growth rate determined for each strain at pH 6.7 is taken as 1.0. (B) Relative OD_{650} after 24 h of *F. prausnitzii* isolates at four bile salt concentrations (0%, 0.1%, 0.25%, and 0.5%) on YCFAG medium. For comparison, the OD_{650} after 24 h of incubation determined for each isolate in medium without bile salt has been taken as 1.0. Mean growth rates at pH 6.7 and mean OD_{650}s in the absence of bile salts for each strain (±standard deviation) were as follows: ●, ATCC 27768 (0.17 ± 0.02 and 0.33 ± 0.05, respectively); ▼, M21/2 (0.32 ± 0.07 and 1.39 ± 0.05, respectively); ■, S3L/3 (0.16 ± 0.02 and 0.52 ± 0.07, respectively); ◆, S4L/4 (0.20 ± 0.02 and 0.63 ± 0.06, respectively); ○, A2-165 (0.55 ± 0.04 and 0.77 ± 0.02, respectively); Δ, L2-6 (0.19 ± 0.01 and 0.47 ± 0.02, respectively); □, HTF-75H (0.15 ± 0.01 and 0.386 ± 0.046, respectively); ◇, HTF-F (0.18 ± 0.01 and 0.826 ± 0.089, respectively). Phylogroup I isolates have been represented in black while phylogroup II isolates are shown in white.

The basal YCFA medium (described in Materials and Methods) contained 33 mM acetate, which is known to stimulate the growth of *F. prausnitzii* strains (11). Growth was assessed where possible by the change in OD_{650}, but for insoluble substrates such as xylan, it was necessary to rely on change in medium pH as an indicator of substrate fermentation. The ability of *F. prausnitzii* to utilize dietary polysaccharides was somewhat limited with no growth on arabinogalactan, no fermentation of xylan, and little or no growth on soluble starch. While two strains grew well on inulin, the remainder grew poorly. Stimulation of *F. prausnitzii* 16S rRNA sequences by inulin has been reported *in vivo* in healthy human volunteers (38), but it appears likely from the present work that this stimulation may favor certain strains. Interestingly, most isolates grew on apple pectin, although not on citrus pectin. Salyers et al. (43, 44) noted that the utilization of uronic acids was unusual in genera from the human colon other than *Bacteroides* species. In the present study, several *F. prausnitzii* strains were able to utilize galacturonic acid, which is an important constituent of pectin.

Growth was also detected for most *F. prausnitzii* strains on the host-derived sugar *N*-acetylglucosamine and for some strains on D-glucosamine and D-glucuronic acid, while β-glucuronidase activity has been reported previously in some *F. prausnitzii* isolates (8). This suggests that *F. prausnitzii* has the ability to switch between diet- and host-derived substrates, in common with several other dominant human colonic species (48). None of the carbohydrates tested allowed differentiation between the two phylogroups.

Very little growth was observed when carbohydrates were omitted from the medium, although the basal YCFA medium contains 1% Casitone. This indicates that *F. prausnitzii* strains have little or no ability to grow with peptides as their sole energy source. No evidence was found for fermentation of porcine gastric mucin.

Tolerance of *Faecalibacterium prausnitzii* isolates to the gut environment. Previous studies have reported that *F. prausnitzii* growth is inhibited by slightly acidic pH (12). The eight isolates tested showed growth rates at pH 5.75 ranging between 20% (for A2-165) and 80% (for HTF-F) of those at pH 6.7 (Fig. 4A). On average, there was a 14% decrease at pH 6.2, but a 60% decrease at pH 5.75, compared with pH 6.7. Tolerance of bile salts, whose concentrations have been reported to increase in certain gut disorders (24, 35), is also considered to be an important factor for survival in the intestine. Bile salt tolerance differed among isolates, particularly at the lowest concentration tested (0.1%), but all the strains tested were bile salt sensitive, showing on average 76%, 95%, and 97% inhibition at 0.1%, 0.25%, and 0.5% bile salts, respectively (Fig. 4B). In contrast, other species of intestinal bacteria such as *Bacteroides* spp. and *Enterococcus faecium* have been reported to be resistant to up to 20% and 40% bile salt concentrations, respectively (5). Bile acids are synthesized in the liver and released into the small intestine, where it is estimated that 90 to 95% of secreted bile is absorbed. The concentration of bile in the healthy large intestine is approximately 0.05 to 0.3%. The sensitivity of all the *F. prausnitzii* isolates tested to bile salts suggests that this is a factor that may restrict populations of this species in regions of high bile concentration, e.g., within the small intestine.

While these differences in sensitivity to bile salts and pH seem likely to influence the distribution of individual strains, there was no statistically significant evidence for consistent differences between phylogroups.

TABLE 3 Distribution of pectin-utilizing ability among cultured strains of human colonic anaerobes

Phylum and species	No. of strains tested[a]	No. of pectin utilizers
Bacteroidetes		
Bacteroides thetaiotaomicron	22	22
Bacteroides ovatus	24	23
Bacteroides vulgatus	22	7
Bacteroides fragilis	53	17
Other *Bacteroides* spp.	67	19
Actinobacteria		
Bifidobacterium spp.	41	0
Collinsella (formerly *Eubacterium*) *aerofaciens*	15	0
Firmicutes		
Eubacterium rectale + *Roseburia* spp.	20; 10[b]	0
Eubacterium eligens	5	3
Eubacterium biforme	5	0
Ruminococcus obeum, *R. torques*, *R. gnavus*	16	0
Coprococcus spp.	7	0
Peptostreptococcus spp.	8	0
Lactobacillus spp.	6	0
Fusobacterium spp.	10	0
Faecalibacterium prausnitzii	10[b]	8[b]
Ruminococcus albus, *R. bromii*, *R. callidus*	14	0
Eubacterium siraeum	2[b]	0
Other (unclassified)	7	0

[a] Unless indicated otherwise, data are from the work of Salyers et al. (43, 44).
[b] This study. One *B. fibrisolvens* strain is included here along with the *E. rectale* plus *Roseburia*-related strains tested, which are detailed in Materials and Methods.

Potential role of *Faecalibacterium prausnitzii* in the fermentation of pectin in the colon. Pectin is extensively fermented in the human colon (7, 55), but the ability to utilize pectin for growth has been reported for relatively few groups of human colonic bacteria. Salyers et al. (43, 44) showed that pectin utilization was relatively common among *Bacteroides* spp., occurring in 47% of 188 isolates surveyed and prompting subsequent studies on *B. thetaiotaomicron* (9, 54). In contrast, of the 154 strains of Gram-positive anaerobes tested, which included five strains reported as *Fusobacterium prausnitzii*, only *Eubacterium eligens* was previously found to utilize pectin or polygalacturonic acid (43) (Table 3). The present data, however, indicate that *F. prausnitzii* could have a major role in pectin utilization (Table 3).

In order to test this hypothesis further, we examined the ability of two *F. prausnitzii* strains (S3L/3 and A2-165) to compete for apple pectin with representatives of the two other known groups of pectin-utilizing bacteria, *B. thetaiotaomicron* and *E. eligens*. As previous studies have shown that pH plays a critical role in determining the outcome of competition between *Bacteroides* spp. and *Firmicutes* (12, 58), incubations were performed at three initial pH values typical of the range seen in the distal colon (Fig. 5; see also Tables S1 and S2 in the supplemental material). In pure cultures, the major fermentation products produced from pectin were butyrate for *F. prausnitzii*, acetate and succinate for *B. thetaiotaomicron*, and formate and acetate for *E. eligens* (Fig. 5A). As previously observed for growth on starch and glucose (12), the lowest pH

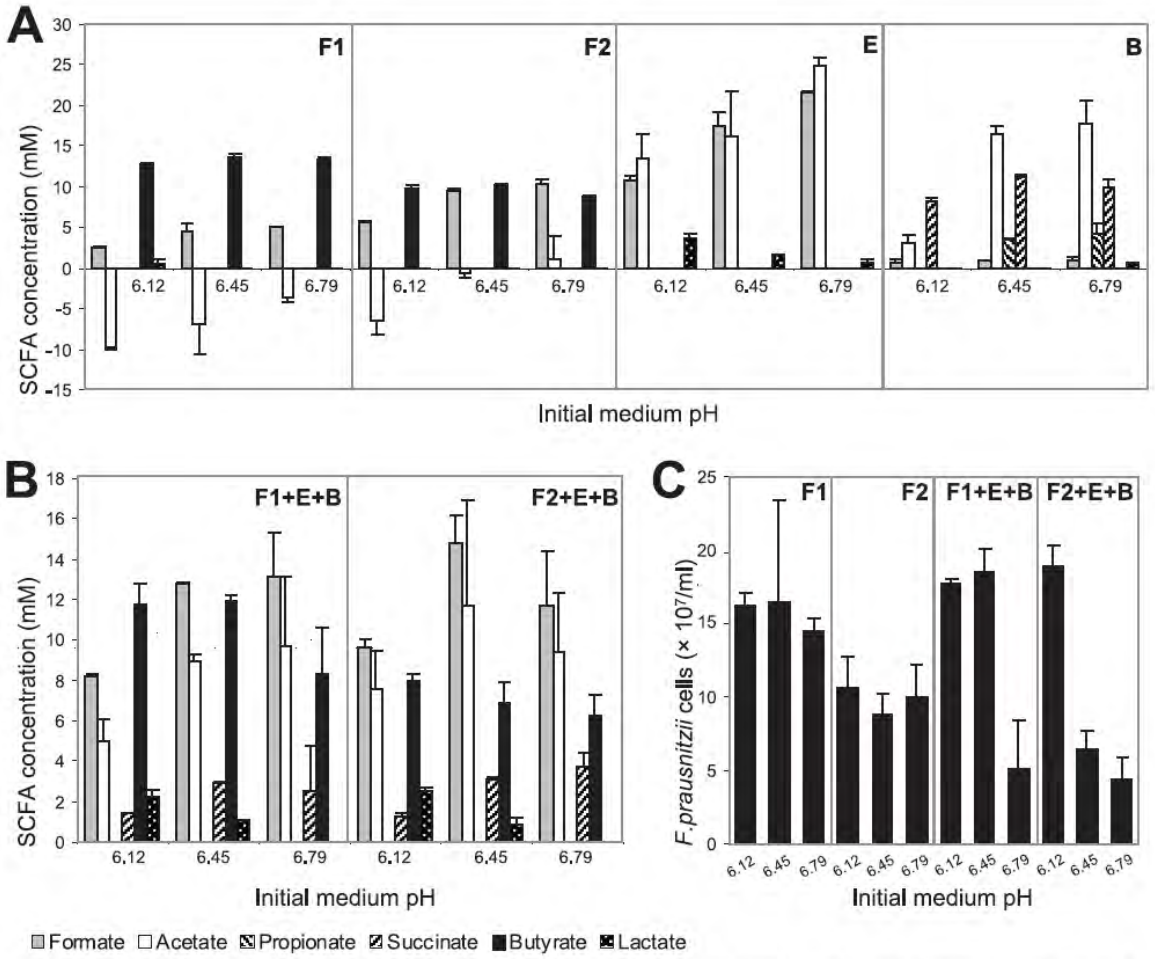


FIG 5 Competition for apple pectin. (A) Change in acidic product concentrations in the growth medium after 24-h fermentation of 0.5% apple pectin by monocultures and cocultures of isolated pectin-utilizing bacteria. F1, *F. prausnitzii* SL3/3; F2, *F. prausnitzii* A2-165; E, *Eubacterium eligens* 3376; B, *B. thetaiotaomicron* 5482 (monocultures). (B) F1+E+B and F2+E+B were tricultures of the three strains indicated. Negative values for acetate reflect the net consumption of acetate initially present in the medium by *F. prausnitzii* strains. Each strain or strain combination was inoculated into YcFA medium adjusted to three different initial pH values (6.12, 6.45, and 6.79). Final medium pH (measured in all cases and detailed in Table S2 in the supplemental material) had decreased after 24 h by up to 0.3 unit for *F. prausnitzii* monocultures, up to 0.7 unit for *B. thetaiotaomicron*, and up to 0.9 unit for *E. eligens*. The final pHs in the tricultures were 6.16 (F1+E+B) and 6.07 (F2+E+B) from initial pH 6.79, 5.79 (F1+E+B) and 5.64 (F2+E+B) from initial pH 6.45, and 5.47 (F1+E+B) and 5.33 (F2+E+B) from initial pH 6.12. Data for two-membered cocultures and on overall sugar utilization from the same experiment are given in Tables S1 and S2 in the supplemental material. (C) Numbers of *F. prausnitzii* cells detected by fluorescent *in situ* hybridization in cultures and cocultures. Counts/ml immediately after inoculation ($t = 0$) were as follows: S3L/3, $0.91 \times 10^7 \pm 0.05 \times 10^7$, and A2-165, $1.31 \times 10^7 \pm 0.01 \times 10^7$.

(6.12) curtailed fermentation of pectin by *B. thetaiotaomicron*. As expected (Fig. 4A), both *F. prausnitzii* strains grew well at the lowest pH (Fig. 5). Tricultures including all three species showed large amounts of butyrate at all three pH values, thus confirming the ability of *F. prausnitzii* to compete for this substrate with the other two pectin-utilizing species (Fig. 5B). Counts estimated by FISH for *F. prausnitzii* after 24 h of incubation indicated greater numbers in the triculture at the lowest pH than at the highest pH (Fig. 5C). Butyrate concentration was less affected by pH, indicating continued fermentative activity by *F. prausnitzii* in spite of decreased cell growth at the highest pH. Data for two-membered cocultures from this experiment are shown in Tables S1 and S2 in the supplemental material. Pectin utilization (measured by decrease in total sugar) was highest for cultures including *B. thetaiotaomicron* at pH 6.79 (see Table S2 in the supplemental material).

Conclusions. *F. prausnitzii* is one of the three most abundant bacterial species found in the healthy adult human large intestine, but its ecology has remained largely unknown. This study has substantially increased the number of cultured, characterized *F. prausnitzii* isolates of human origin and has begun to provide a better understanding of the diversity and microbial ecology of this species in the colon. Based on their 16S rRNA sequences, the available cultured isolates define two broad phylogroups that also include 97% of *F. prausnitzii* 16S rRNA sequences that are detected by direct amplification from human fecal DNA. Our analysis of phylogroup I and II strains from healthy individuals did not reveal systematic differences between the phylogroups with respect to substrate utilization, pH tolerance, or bile sensitivity. Nevertheless, molecular surveys indicate that representatives of the two phylogroups often coexist among the dominant microbiota of individuals (53, 57). There is evidence for reduced representation of

F. prausnitzii in active ileal Crohn's disease (50), and it would be of interest in the future to compare the characteristics, including potential interactions with the immune system, of *F. prausnitzii* strains isolated from CD patients with those from healthy subjects.

Based on our analysis of substrate utilization in 10 cultured strains from seven healthy individuals, most *F. prausnitzii* strains have the ability to utilize apple pectin for growth. The previous report that *F. prausnitzii* strains failed to use pectin is most likely to reflect the use of citrus pectin in that study (43). We have shown that *F. prausnitzii* strains are able to compete for apple pectin as a substrate in the presence of two other known pectin-utilizing species, *B. thetaiotaomicron* and *E. eligens*, suggesting that they make a contribution to pectin fermentation in the colon. Our results suggest that this may apply especially at mildly acidic pH values when competition from *Bacteroides* spp. is reduced (12, 58). The possibility is also raised that certain pectin-rich substrates might be used to develop prebiotic approaches for stimulating *F. prausnitzii* numbers; interestingly, apple pectin has been shown to promote certain *Firmicutes* in a recent study with rats (25). Another notable attribute of some *F. prausnitzii* strains is the utilization of uronic acids for growth, an ability previously thought to be limited to *Bacteroides* spp. among human gut anaerobes. Further analysis of substrate utilization in this species will undoubtedly be aided by the availability of draft genomes for several of the *F. prausnitzii* strains studied here. In conclusion, the present findings demonstrate a broad capacity to utilize both diet- and host-derived growth substrates that helps to explain the remarkable abundance of this species within the human colonic microbiota.

ACKNOWLEDGMENTS

We acknowledge the Scottish Government Rural Environment Research and Analysis Directorate for support. Mireia Lopez-Siles was awarded a fellowship from the Spanish Ministerio de Educación to support a research stay at the RINH and is also the recipient of an FI grant from the Generalitat de Catalunya (2010FI_B2 00135), which receives support from the European Union Commissionate. This work was partially funded by the Spanish Ministry of Education and Science through project SAF2006-0041).

We thank Isabelle Laugaudin for contributing to growth tests, Donna Henderson for SCFA analysis, and Pauline Young for bacterial 16S rRNA sequencing. We also thank Anna Plasencia, Corran Musk, and Carles Borrego for their useful suggestions and technical assistance with phylogenetic analysis and Julien Tap, Petra Louis, and Alan Walker for help in identifying representative sequences for *F. prausnitzii* OTUs from published studies (53, 57) examining 16S rRNA clone libraries.

REFERENCES

1. Balamurugan R, Rajendiran E, George S, Samuel GV, Ramakrishna BS. 2008. Real-time polymerase chain reaction quantification of specific butyrate-producing bacteria, *Desulfovibrio* and *Enterococcus faecalis* in the feces of patients with colorectal cancer. J. Gastroenterol. Hepatol. 23: 1298–1303.
2. Barcenilla A, et al. 2000. Phylogenetic relationships of butyrate-producing bacteria from the human gut. Appl. Environ. Microbiol. 66: 1654–1661.
3. Bryant MP. 1972. Commentary on the Hungate technique for cultivation of anaerobic bacteria. Am. J. Clin. Nutr. 25:1324–1328.
4. Cato EP, Salmon CW, Moore WEC. 1974. *Fusobacterium prausnitzii* (Hauduroy *et al.*) Moore and Holdeman: emended description and designation of neotype strain. Int. J. Syst. Bacteriol. 24:225–229.
5. Cowan S. 1974. Cowan and Steel's manual for the identification of medical bacteria, 2nd ed. Cambridge University Press, London, United Kingdom.
6. Cucchiara S, Iebba V, Conte MP, Schippa S. 2009. The microbiota in inflammatory bowel disease in different age groups. Dig. Dis. 27:252–258.
7. Cummings JH, et al. 1979. The digestion of pectin in the human gut and its effect on calcium absorption and large bowel function. Br. J. Nutr. 41:477–485.
8. Dabek M, McCrae SI, Stevens VJ, Duncan SH, Louis P. 2008. Distribution of beta-glucosidase and beta-glucuronidase activity and of beta-glucuronidase gene gus in human colonic bacteria. FEMS Microbiol. Ecol. 66:487–495.
9. Dongowski G, Lorenz A, Anger H. 2000. Degradation of pectins with different degrees of esterification by *Bacteroides thetaiotaomicron* isolated from human gut flora. Appl. Environ. Microbiol. 66:1321–1327.
10. DuBois M, Gilles KA, Hamilton JK, Rebers PA, Smith F. 1956. Colorimetric method for determination of sugars and related substances. Anal. Chem. 28:350–356.
11. Duncan SH, Hold GL, Harmsen HJ, Stewart CS, Flint HJ. 2002. Growth requirements and fermentation products of *Fusobacterium prausnitzii*, and a proposal to reclassify it as *Faecalibacterium prausnitzii* gen. nov., comb. nov. Int. J. Syst. Evol. Microbiol. 52:2141–2146.
12. Duncan SH, Louis P, Thomson JM, Flint HJ. 2009. The role of pH in determining the species composition of the human colonic microbiota. Environ. Microbiol. 11:2112–2122.
13. Felsenstein J. 2007. PHYLIP (Phylogeny Inference Package) version 3.67. Department of Genetics, University of Washington, Seattle, WA.
14. Flint HJ, Duncan SH, Scott KP, Louis P. 2007. Interactions and competition within the microbial community of the human colon: links between diet and health. Environ. Microbiol. 9:1101–1111.
15. Frank DN, et al. 2007. Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases. Proc. Natl. Acad. Sci. U. S. A. 104:13780–13785.
16. Giovanonni SJ. 1991. The polymerase chain reaction, p 177–201. *In* Stackebrandt E, Goodfellow M (ed), Nucleic acid techniques in bacterial systematics. John Wiley & Sons, New York, NY.
17. Hall TA. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. Ser. 41:95–98.
18. Hamer HM, et al. 2008. Review article: the role of butyrate on colonic function. Aliment. Pharmacol. Ther. 27:104–119.
19. Harmsen HJ, Raangs GC, He T, Degener JE, Welling GW. 2002. Extensive set of 16S rRNA-based probes for detection of bacteria in human feces. Appl. Environ. Microbiol. 68:2982–2990.
20. Hauduroy P, Ehringer G, Urbain A, Guillot G, Magrou J (ed). 1937. Dictionnaire des bactéries pathogenès. Masson and Co., Paris, France.
21. Hold GL, Schwiertz A, Aminov RI, Blaut M, Flint HJ. 2003. Oligonucleotide probes that detect quantitatively significant groups of butyrate-producing bacteria in human feces. Appl. Environ. Microbiol. 69: 4320–4324.
22. Jia W, et al. 2010. Is the abundance of *Faecalibacterium prausnitzii* relevant to Crohn's disease? FEMS Microbiol. Lett. 310:138–144.
23. Lane DJ. 1991. 16S/23S rRNA sequencing, p 115–148. *In* Stackebrandt E, Goodfellow M (ed), Nucleic acid techniques in bacterial systematics. John Wiley and Sons, New York, NY.
24. Lapidus A, Einarsson C. 1998. Bile composition in patients with ileal resection due to Crohn's disease. Inflamm. Bowel Dis. 4:89–94.
25. Licht T, et al. 2010. Effects of apples and specific apple components on the cecal environment of conventional rats: role of apple pectin. BMC Microbiol. 10:13–23.
26. Louis P, et al. 2004. Restricted distribution of the butyrate kinase pathway among butyrate-producing bacteria from the human colon. J. Bacteriol. 186:2099–2106.
27. Louis P, Flint HJ. 2009. Diversity, metabolism and microbial ecology of butyrate-producing bacteria from the human large intestine. FEMS Microbiol. Lett. 294:1–8.
28. Ludwig W, et al. 2004. ARB: a software environment for sequence data. Nucleic Acids Res. 32:1363–1371.
29. Mariat D, et al. 2009. The *Firmicutes/Bacteroidetes* ratio of the human microbiota changes with age. BMC Microbiol. 9:123–128.
30. Martinez-Medina M, Aldeguer X, Gonzalez-Huix F, Acero D, Garcia-Gil LJ. 2006. Abnormal microbiota composition in the ileocolonic mucosa of Crohn's disease patients as revealed by polymerase chain reaction-denaturing gradient gel electrophoresis. Inflamm. Bowel Dis. 12: 1136–1145.
31. Miyazaki K, Martin JC, Marinsek-Logar R, Flint HJ. 1997. Degradation and utilization of xylans by the rumen anaerobe *Prevotella bryantii* (formerly *P. ruminicola* subsp. *brevis*) B(1)4. Anaerobe 3:373–381.

32. Moore WE, Moore LH. 1995. Intestinal floras of populations that have a high risk of colon cancer. Appl. Environ. Microbiol. 61:3202–3207.
33. Muyzer G, de Waal EC, Uitterlinden AG. 1993. Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. Appl. Environ. Microbiol. 59:695–700.
34. Muyzer G, Teske A, Wirsen CO, Jannasch HW. 1995. Phylogenetic relationships of *Thiomicrospira* species and their identification in deep-sea hydrothermal vent samples by denaturing gradient gel electrophoresis of 16S rDNA fragments. Arch. Microbiol. 164:165–172.
35. Pereira SP, Bain IM, Kumar D, Dowling RH. 2003. Bile composition in inflammatory bowel disease: ileal disease and colectomy, but not colitis, induce lithogenic bile. Aliment. Pharmacol. Ther. 17:923–933.
36. Pruesse E, et al. 2007. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. Nucleic Acids Res. 35:7188–7196.
37. Pryde SE, Duncan SH, Hold GL, Stewart CS, Flint HJ. 2002. The microbiology of butyrate formation in the human colon. FEMS Microbiol. Lett. 217:133–139.
38. Ramirez-Farias C, et al. 2009. Effect of inulin on the human gut microbiota: stimulation of *Bifidobacterium adolescentis* and *Faecalibacterium prausnitzii*. Br. J. Nutr. 101:541–550.
39. Richardson AJ, Calder AG, Stewart CS, Smith A. 1989. Simultaneous determination of volatile and non-volatile acidic fermentation products of anaerobes by capillary gas chromatography. Lett. Appl. Microbiol. 9:5–8.
40. Roediger WE. 1980. The colonic epithelium in ulcerative colitis: an energy-deficiency disease? Lancet ii:712–715.
41. Rumney CJ, Duncan SH, Henderson C, Stewart CS. 1995. Isolation and characteristics of a wheat bran-degrading *Butyrivibrio* from human faeces. Letts. Appl. Microbiol. 20:232–236.
42. Saitou N, Nei M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4:406–425.
43. Salyers AA, West SE, Vercellotti JR, Wilkins TD. 1977. Fermentation of mucins and plant polysaccharides by anaerobic bacteria from the human colon. Appl. Environ. Microbiol. 34:529–533.
44. Salyers AA, Vercellotti JR, West SEH, Wilkins TD. 1977. Fermentation of mucin and plant polysaccharides by strains of *Bacteroides* from the human colon. Appl. Environ. Microbiol. 33:319–322.
45. Sartor RB. 2008. Therapeutic correction of bacterial dysbiosis discovered by molecular techniques. Proc. Natl. Acad. Sci. U. S. A. 105:16413–16414.
46. Schloss PD, et al. 2009. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl. Environ. Microbiol. 75:7537–7541.
47. Schwiertz A, et al. 2010. Microbiota in pediatric inflammatory bowel disease. J. Pediatr. 157:240–244.
48. Scott KP, Martin JC, Campbell G, Mayer C-D, Flint HJ. 2006. Whole-genome transcription profiling reveals genes up-regulated by growth on fucose in the human gut bacterium *Roseburia inulinivorans*. J. Bacteriol. 188:4340–4349.
49. Sokol H, et al. 2008. *Faecalibacterium prausnitzii* is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. Proc. Natl. Acad. Sci. U. S. A. 105:16731–16736.
50. Sokol H, et al. 2009. Low counts of *Faecalibacterium prausnitzii* in colitis microbiota. Inflamm. Bowel Dis. 15:1183–1189.
51. Suau A, et al. 1999. Direct analysis of genes encoding 16S rRNA from complex communities reveals many novel molecular species within the human gut. Appl. Environ. Microbiol. 65:4799–4807.
52. Suau A, et al. 2001. *Fusobacterium prausnitzii* and related species represent a dominant group within the human fecal flora. Syst. Appl. Microbiol. 24:139–145.
53. Tap J, et al. 2009. Towards the human intestinal microbiota phylogenetic core. Environ. Microbiol. 11:2574–2584.
54. Tierny Y, Béchet M, Joncquiert JC, Dubourguier HC, Guillaume JB. 1994. Molecular cloning and expression in *Escherichia coli* of genes encoding pectate lyase and pectin methylesterase activities from *Bacteroides thetaiotaomicron*. J. Appl. Bacteriol. 76:592–602.
55. Titgemeyer EC, Bourquin LD, Fahey GC, Garleb KA. 1991. Fermentability of various fiber sources by human fecal bacteria *in vitro*. Am. J. Clin. Nutr. 53:1418–1424.
56. van Tongeren SP, Slaets JP, Harmsen HJ, Welling GW. 2005. Fecal microbiota composition and frailty. Appl. Environ. Microbiol. 71: 6438–6442.
57. Walker AW, et al. 2011. Dominant and diet-responsive groups of bacteria within the human colonic microbiota. ISME J. 5:220–230.
58. Walker AW, Duncan SH, McWilliam Leitch EC, Child MW, Flint HJ. 2005. pH and peptide supply can radically alter bacterial populations and short-chain fatty acid ratios within microbial communities from the colon. Appl. Environ. Microbiol. 71:3692–3700.
59. Wang G, et al. 1993. RAPD (arbitrary primer) PCR is more sensitive than multilocus enzyme electrophoresis for distinguishing related bacterial strains. Nucleic Acids Res. 21:5930–5933.
60. Willing B, et al. 2009. Twin studies reveal specific imbalances in the mucosa-associated microbiota of patients with ileal Crohn's disease. Inflamm. Bowel Dis. 15:653–660.

Chapter 2‡

Mucosa-associated *Faecalibacterium prausnitzii* population profile in healthy and diseased gut

Significance of this study

What was already known on this subject?

- Many studies have evidenced that *F. prausnitzii* is depleted in CD and more recent findings indicate also a reduction in UC. Evidences of depletion of this species have been reported in other disorders such as CRC, IBS, celiac disease and type 2 diabetes.
- Gut environmental conditions are different in healthy and diseased gut. For instance, it has been reported that UC patients have acidic stools, and that the bile salt profile of CD patients is different from that in H subjects.
- *F. prausnitzii* is highly sensitive to oxygen. Growth rate of this species decreases at slightly acidic pH of the culture medium, and the maximum OD of the cultures is compromised by an increase in the bile salt concentration of the medium.

What are the new findings?

- IBD patients host *F. prausnitzii* populations with fewer subtypes than H subjects.
- The main members of *F. prausnitzii* population are shared between H subjects and individuals with gut diseases.
- IBD and CRC *F. prausnitzii* populations can be discriminated from that of H subjects according to the distribution of the common phylotypes and the presence of some disease-specific phylotypes.

How might this impact on research in the foreseeable future?

- This study can serve as basis for depicting the importance of particular subtypes losses in disease pathogenesis.
- Quantification of *F. prausnitzii* phylotypes can be explored as putative biomarkers of disease.

‡

This chapter has been accepted in the journal *Applied and Environmental Microbiology*. Lopez-Siles, M *et al.* **Mucosa-associated *Faecalibacterium prausnitzii* population richness is reduced in inflammatory bowel disease patients.** *Appl. Environ. Microbiol.* 2015; 81 (21): 7582-7592.

Article II

Mucosa-associated *Faecalibacterium prausnitzii* population richness is reduced in inflammatory bowel disease patients

Mireia Lopez-Siles[1], Margarita Martinez-Medina[1], Carles Abellà[1], David Busquets[2], Miriam Sabat-Mir[3], Sylvia H. Duncan[4], Xavier Aldeguer[2], Harry J. Flint[4], and L. Jesús Garcia-Gil[1]✉.

[1] Laboratory of Molecular Microbiology, Biology Department, Universitat de Girona, Girona, Spain. [2] Department of Gastroenterology, Hospital Dr. Josep Trueta, Girona, Spain. [3] Department of Gastroenterology, Hospital Santa Caterina, Salt, Girona, Spain. [4] Microbial Ecology Group, Rowett Institute of Nutrition and Health, University of Aberdeen, Aberdeen, United Kingdom. ✉ Corresponding author, e-mail: jesus.garcia@udg.edu

Applied and Environmental Microbiology, 2015; 81 (21): 7582-7592

Abstract

Faecalibacterium prausnitzii depletion in intestinal diseases has been extensively reported, but little is known about intraspecies variability. This work aims to determine if subjects with gastrointestinal disease host mucosa-associated *F. prausnitzii* populations different from healthy individuals. A new species-specific polymerase chain reaction-denaturing gradient gel electrophoresis (PCR-DGGE) method targeting the 16S rRNA gene was developed to fingerprint *F. prausnitzii* populations in biopsy specimens from 31 healthy control (H) subjects and 36 Crohn's disease (CD), 23 ulcerative colitis (UC), 6 irritable bowel syndrome (IBS), and 22 colorectal cancer (CRC) patients. The richness of *F. prausnitzii* subtypes was lower in inflammatory bowel disease (IBD) patients than in H subjects. The most prevalent operational taxonomic units (OTUs) consisted of four phylotypes (OTUs with a 99% 16S rRNA gene sequence similarity [OTU99]), which were shared by all groups of patients. Their distribution and the presence of some disease-specific *F. prausnitzii* phylotypes allowed us to differentiate the population in IBD and CRC patients from that in H subjects. At the level of a minimum similarity of 97% (OTU97), two phylogroups accounted for 98% of the sequences. Phylogroup I was found in 87% of H subjects but in under 50% of IBD patients (P=0.003). In contrast, phylogroup II was detected in >75% of IBD patients and in only 52% of H subjects (P=0.005). This study reveals that even though the main members of the *F. prausnitzii* population are present in both H subjects and individuals with gut diseases, richness is reduced in the latter and an altered phylotype distribution exists between diseases. This approach may serve as a basis for addressing the suitability of their quantification as putative biomarkers of disease and depicting the importance of the loss of these subtypes in disease pathogenesis.

Introduction

Metagenomic studies have shown that the human gut microbiota is constituted by a relatively limited number of dominating bacterial phyla. While in healthy adults, Bacteroidetes and Firmicutes are the most abundant phyla; Proteobacteria, Verrucomicrobia, Actinobacteria and Fusobacteria are relatively scarce (1-3). The Firmicute *Faecalibacterium prausnitzii* (Ruminococcaceae) is one of the three most abundant species, representing approximately 6 to 8% of the gut microbial community in healthy subjects, although it can reach up to 20% in some individuals (1, 4-11). In contrast, depletion of *F. prausnitzii* has been reported to occur in several pathological disorders (for review see reference (12) and references therein) such as Crohn's disease (CD)(12-19), ulcerative colitis (UC) (11, 14, 15, 17, 20-26), irritable bowel syndrome (IBS) of alternating type (27), colorectal cancer (CRC) (28, 29), and diabetes (30-32).

Many studies have shown the potential role of *F. prausnitzii* in promoting gut health through the secretion of anti-inflammatory compounds such as butyrate (16, 33-36), and in reducing the severity of colitis induced in mice (16, 37). Despite being a relatively abundant bacterium capable of regulating gut homeostasis (38, 39) and interacting in several host pathways (40), relatively few studies have paid attention to the distribution of phylotypes within *Faecalibacterium* populations in the human gut. Random Amplified Polymorphic DNA-polymerase chain reaction (RAPD-PCR) fingerprinting of 18 isolates from faecal samples of 10 healthy subjects showed that an individual can have up to four different *F. prausnitzii* strains and that these are grouped by individual (35). In addition, 16S rRNA gene analysis of these isolates indicated that despite each strain has a unique sequence, but that the isolates group into two *F. prausnitzii* phylogroups that have 97% 16S rRNA gene sequence similarity. These two phylogroups coexist in healthy individuals (35), and comprise approximately 97% of *F. prausnitzii* 16S rRNA sequences found in feces (10, 41). However, it remains to be elucidated how many different *F. prausnitzii* are hosted by patients with gut diseases, and it is still unknown if the *F. prausnitzii* population found in patients suffering from intestinal disorders differs from that found in healthy subjects.

This work describes the *F. prausnitzii* populations present in inflammatory bowel disease (IBD), CRC and IBS patients. The populations were determined by using a species-specific polymerase chain reaction (PCR) followed by denaturing gradient gel electrophoresis (DGGE) and then sequencing of bands. The main objective therefore was to determine if the phylotype profiles correlate with certain intestinal disorders. We also investigated whether or not certain phylotypes are associated with patients' clinical characteristics in order to reveal biomarkers potentially useful for diagnostic support and/or in establishment of a prognosis.

Materials and Methods

Patients, clinical data and sampling

A Spanish cohort consisting of 118 volunteers (36 CD, 23 UC, 6 IBS, 22 CRC, and 31 healthy control [H] subjects) was included (Table 1). Subjects were recruited by the Gastroenterology Services of the Hospital Universitari Dr. Josep Trueta (Girona, Spain) and the Hospital Santa Caterina (Salt, Spain). Patients were gender and age matched, except CRC patients, who were significantly older than all the other groups ($P<0.001$), and H subjects who were older than those with IBD ($P\leq0.013$). IBD was diagnosed according to standard clinical, pathological, and endoscopic criteria and categorized as stated in the Montreal classification (42). Rome III criteria (available at http://www.romecriteria.org/criteria/) were used to diagnose IBS. A CRC diagnosis was established by colonoscopy and biopsy. The control group (H subjects) consisted of subjects with normal colonoscopy findings who underwent this procedure for reasons such as rectorrhagia (N=9), familial history of colorectal cancer (N=11), and abdominal pain (N=11). None of the subjects had received antimicrobial treatment for at least two months before colonoscopy.

Prior to colonoscopy, patients were subjected to cleansing of the gastrointestinal tract using Casenglicol® following the manufacturer's guidelines. During routine colonoscopy, a biopsy sample from the transverse colon was collected from each subject following standard procedures. When it was not technically possible to collect a biopsy sample from the transverse colon, rectal biopsy samples were taken instead, because the mucosa-associated community profile is rather stable along the gut (15, 43). All biopsies specimens were immediately placed in sterile tubes without any buffer and stored at -80 ºC following completion of the whole endoscopic procedure and upon analysis.

Ethical consideration

This work was approved by the Ethics Committee of Clinical Research of the Hospital Universitari Dr. Josep Trueta (Girona, Spain) and the Institut d' Assistència Sanitària of Girona (Salt, Spain) on 24th February 2009 and 21st April 2009, respectively. Informed consent from the subjects was obtained before enrollment.

Table 1. Sample size and clinical characteristics of subjects.

	Healthy*	Irritable bowel syndrome	IBD		Colorectal cancer	p value§
			Ulcerative colitis	Crohn's disease		
N (patients)	31	6	23	36	22	
Age (mean years ± SD)	49.2±16.3	42.4±11.4	38.4±14.0	34.5±12.8	70.1±10.3	<0.001‡
Male (N, %)	15 (48.4%)	2 (20.0%)	15 (65.2%)	21 (58.3%)	11 (50.0%)	0.538†
Active (N, %)	na	na	17 (73.9%)	23 (63.8%)	na	0.365†
Previous surgery (N, %)	0	nd	2 (8.7%)	7 (19.4%)	nd	0.145†
Smokers (N, %)	0	0	2 (8.7%)	4 (11.1%)	1 (4.5%)	0.386†
Treatment (N, %) **						0.520†
No treatment	na	na	10 (43.5%)	11 (30.6%)	na	
Mesalazine	na	na	2 (16.7%)	3 (8.3%)	na	
Moderate immunosuppressant	na	na	4 (17.4%)	11 (30.6%)	na	
Anti-TNFα (infliximab, adalimumab)	na	na	4 (17.4%)	7 (19.4%)	na	
UC classification (N, %) **						na
Ulcerative proctitis (E1)	na	na	4 (17.4%)	na	na	
Distal UC (E2)	na	na	12 (52.2%)	na	na	
Extensive UC or pancolitis (E3)	na	na	4 (17.4%)	na	na	
CD Montreal classification						
Age of diagnosis (N, %) **						0.309‡
diag < 16y (A1)	na	na	1 (4.3%)	3 (8.3%)	nd	
diag 17-40y (A2)	na	na	11 (47.8%)	23 (63.8%)	nd	
diag >41y (A3)	na	na	7 (30.4%)	7 (19.4%)	nd	
Location (N, %)						na
Ileal-CD (L1)	na	na	na	11 (30.5%)	na	
Colonic-CD (L2)	na	na	na	11 (30.5%)	na	
Ileocolonic-CD (L3)	na	na	na	9 (25.0%)	na	
Behavior (N, %) **						na
Non-stricturing, non-penetrating (B1)	na	na	na	20 (55.6%)	na	
Stricturing (B2)	na	na	na	4 (11.1%)	na	
CRC subtype (N, %) **						na
Sporadic	na	na	na	na	11 (50.0%)	
Hereditary***	na	na	na	na	1 (4.5%)	

IBS, Irritable bowel syndrome; IBD, Inflammatory bowel disease; UC, Ulcerative colitis; CD, Crohn's disease; CRC, Colorectal cancer; nd, not determined; na, not applicable

*Controls consisted of subjects who underwent colonoscopy for different reasons: 9/31 rectorrhagia, 11/31 colorectal cancer familial history and 11/31 abdominal pain.

** Medical treatment at the time of sampling was available in 32/36 CD patients, and 20/23 UC patients; Age of disease onset was available for 33/36 CD patients, and 19/23 UC patients; Disease behavior at last follow-up before the time of sampling was available in 24/36 CD patients, and none had penetrating CD (B3); Maximal disease extent at the time of sampling was available in 20/23 UC patients and 31/36 CD patients; presence or absence of relatives with CRC could only be clearly tracked in 12/22 CRC patients.

***Patients were included within this category if a first grade relative has had also CRC.

§ Groups were compared by non-parametric statistical tests, and p value ≤0.05 was considered significant, † χ^2 test, ‡ Mann-Whitney U test

Sample treatment and DNA extraction

Prior to DNA extraction, the biopsy specimens were subjected to two mild ultrasound wash cycles as previously reported (15) to discard transient and loosely attached bacteria. DNA was extracted using the NucleoSpin® Tissue Kit (Macherey-Nagel GmbH &Co., Germany). The support protocol for Gram-positive bacteria and the RNAse treatment step were carried out. Genomic DNA was stored at -80 ºC until use. DNA concentration and purity of the extracts were determined with a NanoDrop ND-100 spectrophotometer (NanoDrop Technologies, USA).

Design of primers to study *F. prausnitzii* populations

A conventional PCR assay consisting of a species-specific primer set targeting the 16S rRNA gene was designed. 16S rRNA gene sequences from *F. prausnitzii* and from other Ruminococcaceae (see Table S1 in the supplemental material) were recovered from GenBank and aligned using the Clustal W program (44). Specific primers targeting DNA regions exclusive to *F. prausnitzii* were manually designed and further checked using Primer Express® (version 3.0) software (Applied Biosystems, Foster City, CA, USA) and NetPrimer® software (available at http://premierbiosoft.com/netprimer, PREMIER Biosoft International, California, USA) to check for primer-dimer structures, hairpins and possible cross dimer interactions between oligonucleotides.

The final primer set designed in this study consisted of primers Fpra 427F (5'-TAAACTCCTGTTGTTGAGGAAGAT-3') and Fpra 1127R (5'-TTTGTCAACGGCAGTCYKG-3'), whose sequence flank an ~700-bp fragment that includes variable regions V3 to V6.

Oligonucleotides specificity was tested *in silico* by comparing the sequences with those in the Ribosomal Database Project II (45) and GenBank database through the use of Seq Match and BLAST (46) tools, respectively. Additionally, an *in vitro* inclusivity-exclusivity test was performed (see Table S2 in the supplemental materials for details). Primer set coverage was evaluated using the SILVA TestPrime (version 1.0) program (available at http://www.arb-silva.de/search/testprime/). A sensitivity test to determine the detection limit of the reaction (i.e., the lowest concentration at which 95% of the positive samples were detected (47)) was performed (see text in the supplemental material).

PCR amplification and DGGE fingerprinting

For *F. prausnitzii* population profiling, the 16S rRNA gene was partially amplified from extracted genomic DNA using the new primer set designed in this study. To obtain PCR products suitable for DGGE analysis, a GC-rich sequence was attached to the 5' end of the forward primer (5'- CGCCCGCCGCGCCCCGCGCCCGGCCCGCCGCCCCCGCC-CC-3') as described by Muyzer et al. (48). All primers were obtained from Metabion (Steinkirchen, Germany).

PCR reaction was optimized by testing different concentrations of $MgCl_2$ (0.25 to 15 mM), deoxyribonucleoside triphosphates (dNTPs; 0.1 to 3 mM), and primers (50 to 900 nM of each primer). The optimized reaction mixture was used in all samples and was composed of: 1× of buffer (II) (10×; Applied Biosystems, Foster City, California), 2.5 mM of $MgCl_2$ (25 mM; Applied Biosystems), 0.2 mM of dNTPs (10 mM; Applied Biosystems), 0.2 μM of GC-Fpra 427F primer, 0.2 μM of Fpra 1127R (10 pmol/mL each), 0.05 U/μL of AmpliTaq DNA polymerase (5 U/mL, Applied Biosystems), and 1 μL of genomic DNA as a template in a total volume of 50 μL.

All PCR were performed in a GeneAmp PCR System 2700 thermocycler (Applied Biosystems, Perkin-Elmer, CA, USA). PCR conditions were optimized by testing different annealing temperatures (52ºC to 70ºC). The optimized cycling program was used and consisted of 10 minutes at 95ºC for initial denaturation and DNA polymerase activation, followed by 35 cycles of 30 s at 95ºC (denaturation), 30 s at 65ºC (annealing), and 1 min at 72ºC (extension) and then a final extension of 10 minutes at 72ºC. The products were visualized under UV light after gel electrophoresis on 1.5% (wt/vol) agarose gels in 1× TAE buffer (Tris-acetate-EDTA, pH 8.0) stained with ethidium bromide (0.5 μg/mL).

DGGE was carried out with an Ingeny-phorU2 system (Ingeny, Goes, The Netherlands) in 6% (w/v) acrylamide gels with a vertical denaturing gradient ranging from 30% to 70% urea-formamide. Electrophoreses were run in 0.5× TAE buffer at 60ºC and at a constant voltage of 120 V for a minimum of 16 h. Gels were then stained with 1× SYBR gold (Molecular Probes Europe, Invitrogen) for 45 minutes in the dark, visualized under UV light, and photographed.

Sequencing, sequence editing and analyses

All detectable bands were excised from the gel and DNA was extracted as previously reported (15). The DNA was then reamplified by PCR as described above, except that the forward primer Fpra 427F without the GC clamp was used. Positive PCR products were cleaned and sequenced in both directions (forward and reverse) by Macrogen Inc. (Seoul,

Korea). Prior to analysis, sequence chromatograms were manually inspected and none presented double peaks. The quality of the sequences was also checked with Sequence Scaner Software (version 2; Applied Biosystems, Foster City, CA, USA). Only those with high trace scores (the trace score is the average basecall quality value of the bases in the sequence after it is trimmed) were considered for further analyses. Sequences were assembled to obtain high-quality consensus sequences, which were further manually refined using BioEdit sequence alignment editor version (version 7.0.9.0) (49). The presence of chimeras was subsequently checked using DECIPHER (50).

To validate the accuracy of the sequences, two biopsies from the same patient were analyzed as previously described, and identical consensus sequences were obtained from the bands recovered (data not shown).

Consensus sequences were compared against the sequences in the NCBI database by using the BLASTN search tool (46) in order to determine the closest previously reported sequence and isolate. An alignment of the consensus sequences was performed using the Clustal W program (44), with manual curation if needed. For further analysis, sequences of 581 nucleotides in length (the region from V3-V6, which consists of positions 525 to 1106 in the numbering for the *Escherichia coli* sequence and which corresponded to those positions recovered for all sequences) were used.

A neighbor-joining (NJ) (51) distance matrix using the Jukes-Cantor (JC) correction was calculated in Mothur (http://www.mothur.org) (52), which was then used to assign sequences to Operational Taxonomic Units (OTUs) using the farthest neighbor method at cutoffs of 0.03 and 0.01 (i.e., minimum similarities of 97% [OTU97] and 99% [OTU99] were required for any pair of sequences to belong to the same phylogroup or phylotype, respectively). Representative sequences for each OTU (hereafter named OTU97 and OTU99 sequences, respectively) were identified and used for further analyses of the distribution by group of patients. We refer to the OTUs defined at >99% sequence identity as "phylotypes" and those defined at >97% sequence identity as "phylogroups". Unique sequences (100% similarity, hereafter named OTU100 sequences) were also considered to compare sequence-based population composition and to study the richness by groups of patients.

Statistics

The clustering of all samples was performed according to the presence or absence of unique sequences by hierarchical analysis by the intergroup joining method based on the Dice coefficient. This same analysis was performed for groups of patients clustering according to their OTU99 and OTU97 sequence compositions.

The nonparametric Kruskal–Wallis and the Mann–Whitney U tests were used to test for differences in richness for variables with more than two categories and for pairwise comparisons, respectively, according to subject diagnostics, IBD subtype, activity status (active CD and UC patients when Crohn's disease activity index was > 150 (53) and the Mayo score was >3, respectively), smoking habit, intestinal resection, treatment received and tumor state (T1 to T4) for CRC patients.

Pearson's X^2 test was carried out to compare the prevalence of OTUs defined at a 99% and 97% similarity, taking into account the same clinical variables mentioned above in order to determine OTUs specific for each condition.

All statistical analyses were conducted via SPSS software (version 15.0; (SPSS Inc., Chicago, IL). Significance was established when the P value was ≤0.05.

Nucleotide sequence accession numbers

The 16S rRNA gene sequences obtained in this study were deposited in the GenBank/EMBL/DDBJ database under the accession numbers KP005458 through KP005741.

Results

Features of the novel PCR-DGGE assay to study *F. prausnitzii* population

In this study, a novel oligonucleotide set was designed to specifically amplify *F. prausnitzii* (see information in the supplemental material). The detection limit of the reaction was 2,623 target genes, which indicates that if we had many replicate samples with 2,623 copies of 16S rRNA gene of *F. prausnitzii*, no more than 5% failed reactions should occur. The banding pattern obtained by DGGE represents the major constituents of the community analyzed (54). When universal primers are used, species that contribute <1% of the total population would not be readily detected by this molecular approach (48). This is expected to have a minor effect in the present study, however, because *F. prausnitzii* represents at least 6% of the mucosa-associated microbial community in healthy subjects (1, 4-11), and species-specific primers have been used.

The *in silico* analysis of the oligonucleotide set chosen showed that primer Fpra427F was specific for *F. prausnitzii* and targeted all the isolates, whereas primer Fpra1127R was genus-specific. The coverage provided by the Fpra 427F-Fpra1127R primer set was of 70.6% of the *Faecalibacterium* sequences in the SILVA dataset. The remaining 29.4% should be regarded as sequences of this genus but not sequences of *F. prausnitzii* (other species that have not yet been described probably exist within this genus).

Inclusivity-exclusivity tests were conducted *in vitro* by testing DNA from nine *F. prausnitzii* isolates and from 71 additional representative bacterial species (see Table S2 in the supplemental materials), and these confirmed that the PCR was totally specific. All the *F. prausnitzii* isolates were successfully detected, thus producing a single PCR product, and there was no cross-reaction with any of the nontarget microorganisms (see the supplemental materials).

Sequence analysis and *F. prausnitzii* population composition

From the samples from the 118 volunteers engaged in the study, a total of 284 partial *F. prausnitzii* 16S rRNA gene sequences (66 from H, 17 from IBS, 48 from UC, 97 from CD, and 56 from CRC) were obtained from PCR-DGGE gels. This represented 88.75% of the observed bands, whereas the remaining (i.e. 11.25% of the observed bands) were not recovered due to methodological problems (i.e. improper band excision from the gel, an incapability to reamplify the bands for further sequencing) and thus were not included in the analysis. Among the sequences retrieved, 135 unique sequences were found, and these could be grouped into 40 OTUs on the basis of the >99% sequence similarity criterion and 5 OTUs on the basis of >97% of sequence similarity criterion (see Tables S3 to S5 in the supplemental material). The correspondence between OTUs at different cutoffs of similarity is shown in Table S6 in the supplemental material (i.e. which OTU100 sequences are included within a given OTU97 and OTU99 sequences).

F. prausnitzii population similarity and richness by diagnostics (OTU100 sequence analysis)

The sequences were grouped in 135 unique sequences (i.e., OTU100; see Table S3 in the supplemental material), and analyzed in order to compare sequence-based population composition and to estimate *F. prausnitzii* richness between groups of patients.

Clustering analysis of the OTU100 sequences hosted by each subject revealed that the *F. prausnitzii* populations were rather individual specific, as few patients featured identical population composition. Furthermore, 24.6% of the subjects had an *F. prausnitzii* population composition very different from that found in any other patient studied, thus featuring their own branching point in the cluster analysis (Fig. 1). Noticeably, these were mainly H subjects (54.8%), whereas only 16.7% of IBS patients, 21.7% of UC patients, 4.5% of CRC patients, and 22.2% of CD patients had particular *F. prausnitzii* population profiles composed of a set of sequences not found in any other patient (P=0.001). Within the group of CD patients, over 54% of the patients with colonic CD featured unique *F. prausnitzii* populations, but unique *F. prausnitzii* populations were found in approximately

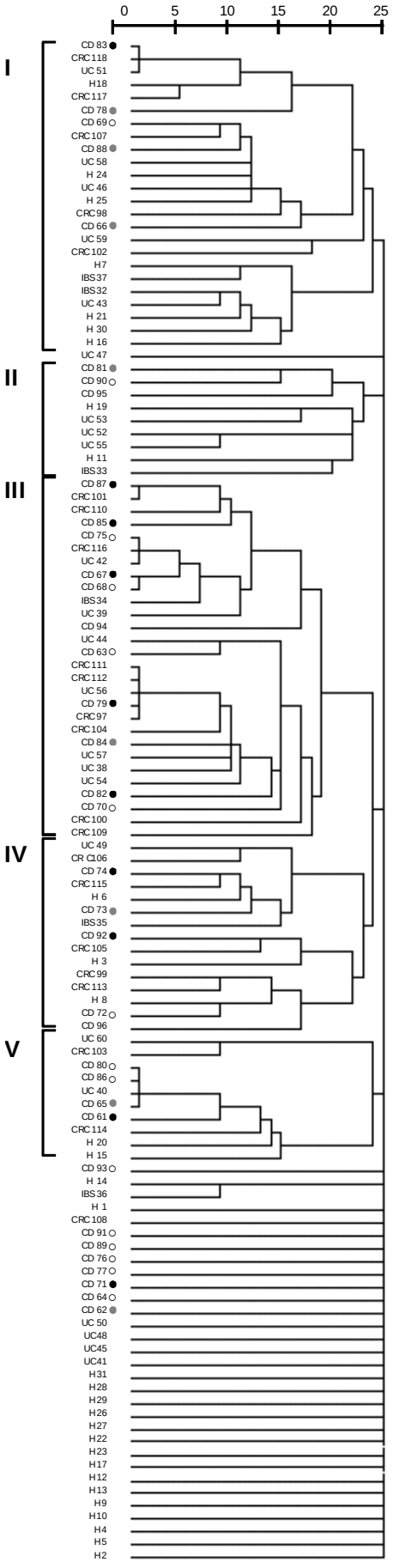


Fig 1. Hierarchical distance clustering showing sample relationship based on *F. prausnitzii* population as determined by comparing unique OTU100 sequences (using the inter-groups joining method based on the Dice coefficient). Scale bar describes similarity between profiles. H, indicates healthy subject; IBS, irritable bowel syndrome; UC, ulcerative colitis; CRC, colorectal cancer; CD, Crohn's disease; ○, colonic-CD (C-CD); ●, ileocolonic-CD (IC-CD); and ●, ileal-CD (I-CD). I to V indicate the main groups in which patients have clustered.

10% of patients with ileal disease location (either ileal CD or ileocolonic CD) (P=0.025).

The sequences from the remaining 75.4% of the subjects grouped into five main clusters. Interestingly, sequences from none of the H subjects were grouped in cluster III, which included the sequences from over 30% of IBD and CRC patients (P=0.005). This clustering was not explained by any other patient variable tested (age, gender, smoking habit, disease activity index, age at the onset of disease, intestinal resection, and medication treatment).

With regard to *F. prausnitzii* richness, all volunteers had populations comprising from one to five unique sequences of this species. The average number of unique sequences of *F. prausnitzii* per subject was lower in IBD patients (1.7±0.8 in UC patients and 1.6±0.8 in CD patients) than in H, IBS and CRC subjects (2.1±1.1 in H subjects, 2.2±1.0 in IBS patients and 2.5±1.2 in CRC patients) (P=0.064). The clinical data for the patients did not explain differences in the number of unique sequences found.

The percentage of subjects with three or more unique *F. prausnitzii* sequences was higher in the H, IBS and CRC groups in comparison with IBD patients (P=0.027) (Fig. 2). These data suggest that IBD patients featured less *F. prausnitzii* microdiversity, whereas CRC and IBS patients present a higher number of unique *F. prausnitzii* subtypes in the gut, with the numbers being similar to the numbers found in H subjects.

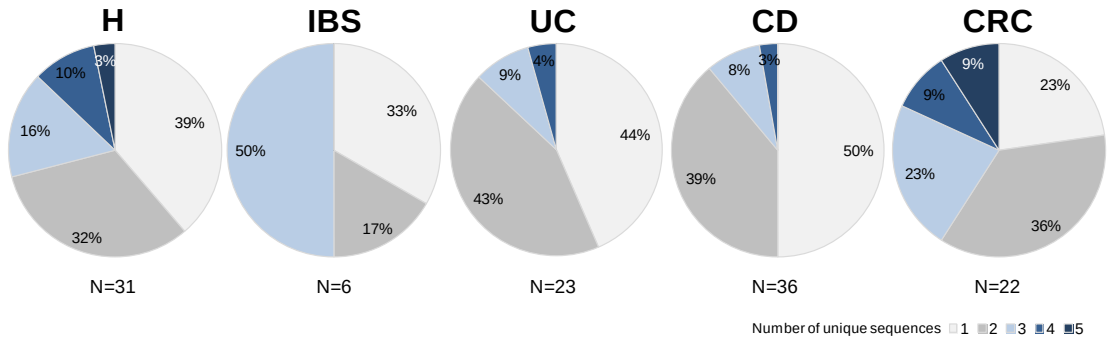


Fig 2. Percentage of patients with *F. prausnitzii* populations with one to five unique sequences (OTU100) by group of patients. H, healthy subject; IBS, irritable bowel syndrome; UC, ulcerative colitis; CD, Crohn's disease; CRC, colorectal cancer.

F. prausnitzii population composition at the phylotype level (OTU99 sequence analysis)

For analysis of the *F. prausnitzii* population composition at the phylotype level, all the samples from patients with the same diagnosis were analyzed together in order to compare the *F. prausnitzii* populations hosted by H subjects and patients with intestinal disorders. We

refer to the 40 OTUs defined according to >99% sequence identity as "phylotypes". All volunteers had populations with from one to four phylotypes of this species, but no significant differences in the average number of OTU99 sequences per patient were observed between groups (P=0.558). In H subjects 20 different phylotypes were recovered; however in IBS, UC, CD and CRC patients, the values were lower (7, 9, 16, and 14 respectively) (Fig. 3). The highest complexity of the community found in H subjects was confirmed from estimates of the Shannon diversity index (H'_H=2.24), whereas the lowest diversity index was calculated for CD patients (H'_{CD}=1.90), even though more subjects in the cohort analyzed had this disorder. Patients with other gut disorders also presented lower diversity index values than H subjects (H'_{IBS}=1.49, H'_{UC}=1.58, H'_{CRC}=1.83).

Cluster analysis of phylotypes by patient group revealed that those with IBD and CRC host *F. prausnitzii* populations different from those hosted by H subjects (Fig. 3). The most prevalent phylotype (OTU1 among the OTU99 sequences [OTU99_1]) accounted for approximately 20% of the sequences in H and IBS subjects and about 40% in IBD and CRC patients (P=0.002) (Fig. 3: see also Table S4 in the supplemental material). The next three most prevalent OTUs (OTU99_2, OTU99_3, and OTU99_4), were also detected in all the patient groups (Figure 3; see also Table S4 in the supplemental material), representing between 8% and 50% of the sequences, depending on the patient group, but the differences in their prevalence between groups of patients did not reach statistical significance. Twelve OTUs (representing 54.8% of the sequences recovered from H subjects) were exclusive of the H group. On the other hand, 22.2% of sequences from CD patients (seven OTUs) were not shared with patients with other intestinal disorders. Similarly, four UC patient-specific OTUs and four CRC patient-specific OTUs were also found, accounting for the 17.4% and the 18.2% of the sequences from each group of patients, respectively. Finally, OTU99_17 accounted for 16.7% of sequences from IBS patients and was only found in this group of patients.

Altogether these observations suggest that the most prevalent *F. prausnitzii* phylotypes (OTU99_1 through OTU99_4) are present in both H subjects and patients with disease, but that rare phylotypes (OTU99_5 to OTU99_40) not found in all the patients exist and could have emerged or disappeared under certain gut conditions, since they are disease specific or exclusively found in healthy subjects.

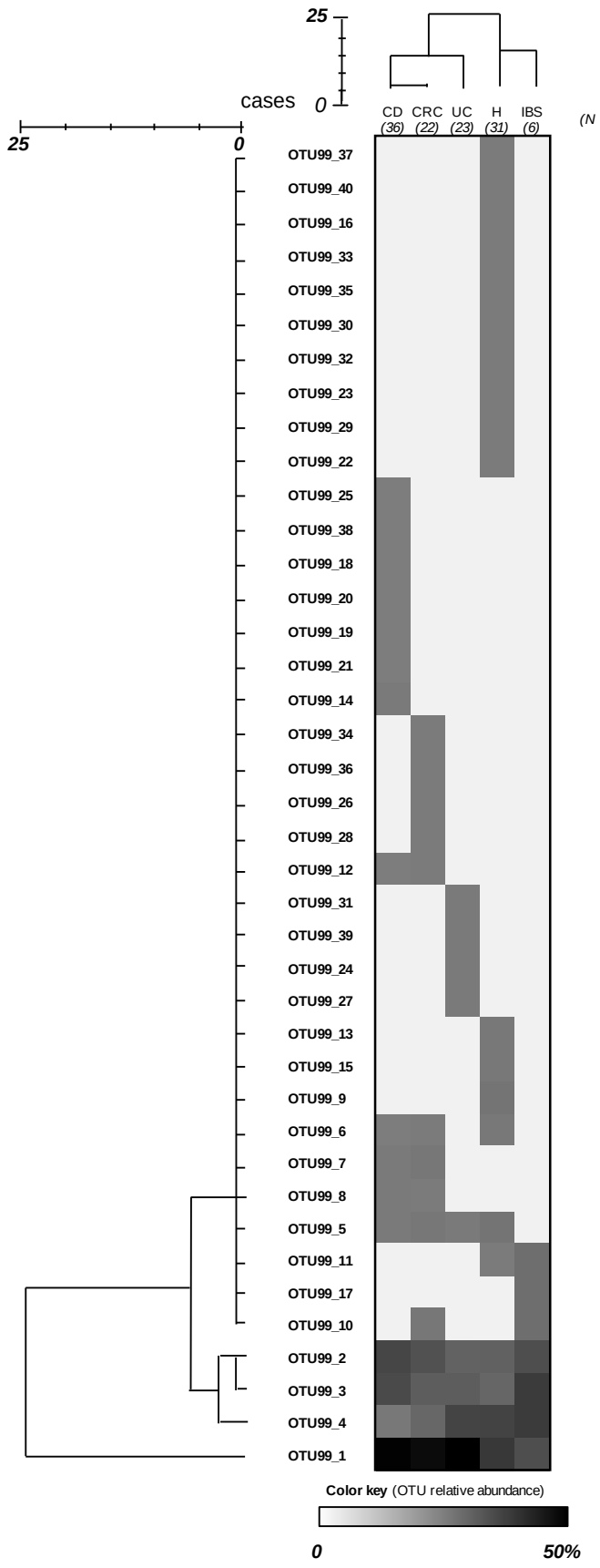


Fig 3. Heat map showing the relative abundance of sequences assigned to each operational taxonomic unit at a 99% 16S rRNA gene sequence similarity (OTU 99) by group of subjects. Relative abundance has been calculated as the percentage of sequences in an OTU from the total of sequences recovered in that group of patients. Hierarchical distance clustering showing relationship of OTU 99 and groups of patients based on their distribution has been represented (using the inter-groups joining method based on the Dice coefficient). Scale bar describes similarity between profiles.
H, healthy subject; IBS, irritable bowel syndrome; UC, ulcerative colitis; CD, Crohn's disease; CRC, colorectal cancer.

F. prausnitzii phylogroup distribution by diagnostics (OTU97 sequence analysis)

Of the total 284 sequences recovered, 279 sequences grouped together in two main groups by OTU97 analysis, and these corresponded to the previously defined phylogroups I and II (35) (Fig. 4; see also Table S5 in the supplemental material). Thus, we refer to the OTUs defined according to a >97% sequence identity as "phylogroups". Phylogroup II (OTU1 among the OTU97 sequences [OTU97_1]) consisted of 56.7% of the total sequences recovered (corresponding to 161 sequences), whereas phylogroup I (OTU97_2) was slightly less prevalent (41.55%, 118 sequences). The remaining five sequences, representing less than 2% of the sequences recovered, were grouped into three different OTUs: a tripleton with sequences from two CD patients and one CRC patient (OTU97_3), a singleton with a sequence from an IBS patient (OTU97_4), and a singleton with a sequence from an H subject (OTU97_5). Noticeably, OTU97_5 presented 95% of similarity to the nearest sequence present in databases (see Table S5 in the supplemental material).

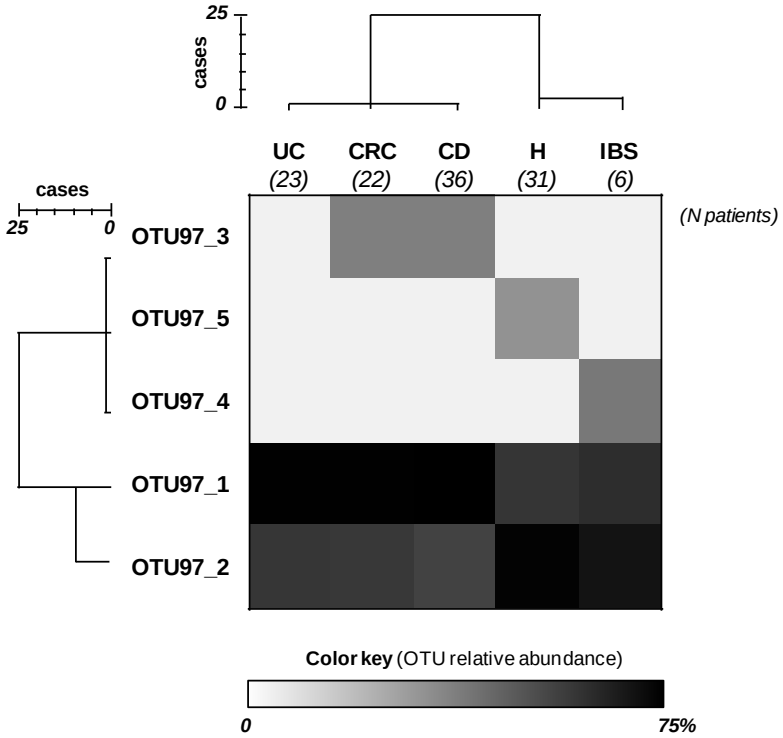


Fig 4. Heat map showing the relative abundance of sequences assigned to each operational taxonomic unit at a 97% 16S rRNA gene sequence similarity (OTU 97) by group of subjects. Relative abundance has been calculated as the percentage of sequences in an OTU from the total of sequences recovered in that group of patients. Hierarchical distance clustering showing relationship of OTU 97 and groups of patients based on their distribution has been represented (using the inter-groups joining method based on the Dice coefficient). Scale bar describes similarity between profiles.
H, healthy subject; IBS, irritable bowel syndrome; UC, ulcerative colitis; CD, Crohn's disease; CRC, colorectal cancer.

All subjects had populations with from one to three phylogroups of *F. prausnitzii*. Significant differences in the average number of OTUs per patient were not found between groups by OTU97 sequence analysis (P= 0.285). However, the prevalences of phylogroups I and II prevalences differed between groups of patients (Fig. 4). More than 52% of the sequences from H and IBS subjects belonged to phylogroup I, whereas most of the sequences from UC, CD and CRC patients belonged to phylogroup II (62.5%, 66%, and 62.5%, respectively) (P=0.001), suggesting that phylogroup I is more frequently compromised in the latter. Cluster analysis of phylogroups by patient group also revealed that those with IBD and CRC host different *F. prausnitzii* populations than H subjects at this cutoff level (Fig. 4).

Analysis of the co-occurrence of both phylogroups (Fig. 5) showed that only 26.1% of UC patients and 22.2% of CD patients harbored both phylogroups simultaneously, while more than 38% of H, IBS and CRC subjects had sequences from both phylogroups, although differences did not reach statistical significance (P=0.270). It is of note that no ileal CD patients had simultaneously both phylogroups (P=0.060).

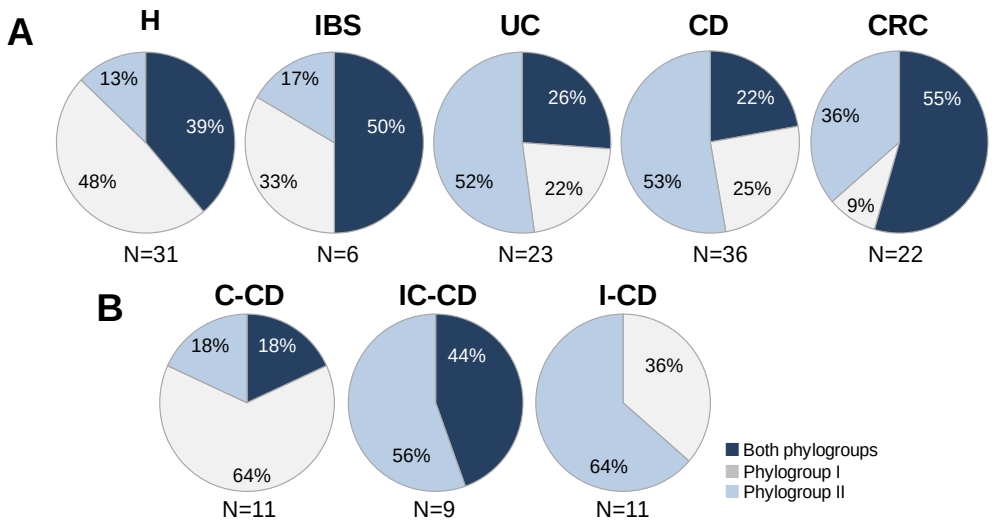


Fig 5. Prevalence of *F. prausnitzii* phylogroups (OTU97) in healthy subjects and patients with different intestinal disorders (A) and by Crohn's disease subtype (B). H, healthy subject; IBS, irritable bowel syndrome; UC, ulcerative colitis; CD, Crohn's disease; CRC, colorectal cancer; C-CD, colonic-CD; IC-CD, ileocolonic-CD; I-CD, ileal-CD.

These results suggest that patients with intestinal disorders feature an altered prevalence of phylogroups, mostly characterized by the presence of monophylogroup populations in some IBD patients, especially those with ileal CD.

Putative indicator sequences for differential diagnosis and/or disease prognostics

We observed differences in the prevalence of phylogroups and phylotypes between groups of patients. Therefore, we further explored which OTUs could be considered potential biomarkers for disease diagnosis. The prevalence of four OTUs by OTU99 sequence analysis was statistically significantly different between the groups of patients ($P \leq 0.027$) (Table 2). In particular, OTU99_1 was found to be more prevalent in IBD and CRC patients than in H subjects ($P=0.020$), OTU99_10 was found in 16.7% of IBS patients and 9.1% CRC patients but not in H nor in IBD subjects ($P=0.027$), and OTU99_11 and OTU99_17 were exclusively found in a 16.7% of IBS patients ($P=0.001$).

Several OTUs were found by OTU99 analysis to be associated with the clinical characteristics of the particular disease. Within the group of IBD patients, OTU99_24 was exclusively found in a 33% of patients with inactive UC ($P=0.015$), whereas OTU99_8 was especially absent in patients with active CD ($P=0.008$). Remarkably, OTU99_8 and OTU99_4 were found in all CRC patients featuring the most severe tumor state (T4) ($P<0.001$ and $P=0.033$ respectively). Moreover, OTU99_7 was not found in CRC patients with tumor states T3 and T4, but it was carried by all CRC patients with T1 tumor state, 12.5% of CRC patients with T2 tumor state, and a 5.6% of CD patients ($P=0.013$).

Referring to phylogroups (OTUs by OTU97 analysis), 87.1% of H subjects and 83.3% of IBS patients had phylogroup I, whereas the proportions were reduced to 63.6% for CRC, 47.8% for UC patients, and 43.2% for CD patients ($P=0.003$) (Table 2). In contrast, a higher prevalence of phylogroup II was observed in IBD and CRC patients (78.3% in UC patients, 83.8% in CD patients, and 90.9% in CRC patients) than in H and IBS subjects (51.6% and 66.7%, respectively) ($P=0.005$). No further differences in relation to patients' clinical characteristics were observed at the phylogroup level.

Discussion

In this study we describe for the first time the richness and the distribution of phylotypes and phylogroups of *F. prausnitzii*. Our data reveal that *F. prausnitzii* populations are rather individual specific, especially in H subjects, in line with previous evidences from studies of *F. prausnitzii* isolates, which were found to group by individual according to their 16S rRNA gene sequence and their RAPD-PCR fingerprint (35). Despite the fact that we have observed that the *F. prausnitzii* populations can include up to five different sequences per individual, we found that in over 87% of IBD patients their *F. prausnitzii* populations consisted of just one or two different sequences. In contrast, at least 30% of H, IBS and CRC subjects hosted populations with more than three different sequences. Multiple

Table 2. Main differences observed at different 16S rRNA gene sequence similarity cutoffs after comparisons of sequences from healthy subjects and patients with several gut disorders.

Similarity cutoff*	OTU	Total sequences	Nearest *F. prausnitzii* isolate			Number of patients (%)					p-value
			Strain	Similarity (%)	Accession Number	H (n=31)	IBS (n=6)	UC (n=23)	CD (n=36)	CRC (n=22)	
99%	OTU99_1	104	HTF-I	99	HQ457031.1	12 (38.7)	2 (33.3)	17 (73.9)	24 (66.7)	16 (72.7)	**0.020**
	OTU99_10	3	A2-165	99	AJ270469.2	0 (0)	1 (16.7)	0 (0)	0 (0)	2 (9.1)	**0.027**
	OTU99_11	2	S4L/4	99	HQ457025.1	0 (0)	1 (16.7)	0 (0)	0 (0)	0 (0)	**0.001**
	OTU99_17	1	A2-165	97	AJ270469.2	0 (0)	1 (16.7)	0 (0)	0 (0)	0 (0)	**0.001**
97%	OTU97_1	161	HTF-I	99%	HQ457031.1	16 (51.6)	4 (66.7)	18 (78.3)	31 (86.1)	20 (90.9)	**0.005**
	OTU97_2	118	S3L/3	99%	HQ457024.1	27 (87.1)	5 (83.3)	11 (47.8)	16 (44.4)	14 (63.6)	**0.003**
	OTU97_4	1	A2-165	97%	AJ270469.2	0	1(16.7)	0	0	0	**0.001**

H, healthy subject; IBS, irritable bowel syndrome; UC, ulcerative colitis; CD, Crohn's disease; CRC, colorectal cancer.
* Different phylotypes found as calculated by Mothur with the farthest neighbor method using a cutoff of 99% or 97% similarity of the 16S rRNA gene sequence.

16S rRNA gene copies, with slight sequence variation can be hosted by a bacterial species, which can affect bacterial community analyses (55). According to data from the *rrn*DB database (56), *F. prausnitzii* S3L/3 has a single copy of the 16S rRNA gene, which makes our results more meaningful. To date, only one genome of this species has been fully sequenced and annotated; therefore, the possibility that future studies will reveal that several copies of the 16S rRNA gene are hosted by other members of this species cannot be ruled out. This will reduce the number of unique phylotypes observed per patient. In our study, no differences in the average number of unique sequences per patient were observed according to whether the patient had active or inactive IBD, the IBD or CRC subtype, intestinal resection, use of medication, or smoking habit, suggesting that richness remains reduced over time even if there is endoscopic and clinical signs of remission and regardless of the treatment used.

BLAST analysis revealed that the sequences recovered were of high similarity to the sequences of previously characterized strains of this species. Only two of the sequences featured ≤95% similarity with the *F. prausnitzii* sequences found previously. This finding suggests that novel and rare phylotypes of *F. prausnitzii* are yet to be retrieved by cultivation techniques. The sequences of both phylogroups I and II (35) were detected in all groups of patients; but in some patients, especially IBD patients, we detected only one of the two main phylogroups. In general, all patients suffering from an intestinal disorder exhibited a reduction in phylotype richness which was not recovered during periods of remission or in patients with mild states of the disease, suggesting that alterations in this population struggle to normalize with the patient's current treatments. New therapies to recover all the diversity of *F. prausnitzii* in these patients should be considered.

Differences in phylotype and phylogroup prevalence between patients with different diagnoses allowed us to discriminate patients suffering from intestinal disease, especially those with IBD and CRC, from H subjects. This finding is in agreement with the findings of previous studies reporting that phylotypes related to isolate M21/2 (phylogroup I) and/or isolate A2-165 (phylogroup II) are depleted in CD patients compared to H subjects (15, 22). As these results might indicate differences in the abundance of these phylogroups, it would be of interest to conduct quantitative analyses in order to explore their usefulness as biomarkers for the diagnosis and/or prognosis of intestinal disorders. Additionally, further investigation addressing the role of disease-specific sequences and the absence of H subject-associated phylotypes in patients with gut disease could shed light on how *F. prausnitzii* can contribute to or prevent pathogenesis of gut diseases.

Our data have revealed that H subjects and those with intestinal disorders host distinct *F. prausnitzii* populations. These distinct populations can mainly be attributable not

only to differences in the prevalence of the common *F. prausnitzii* phylotypes but also to the presence of rare OTUs found specifically in each group of patients. These differences in the presence or absence of a specific phylotype can be explained by different abilities to adapt to a diseased gut environment or mucosal status. For instance, it has been reported that CD patients often have acidic stools with elevated bile salt concentrations (57, 58), and other factors, such as oxidative stress or thiol availability, are also likely to lead to a significantly altered microbiota (59, 60). Supporting this hypothesis, *F. prausnitzii* representatives are sensitive to changes in the gut physicochemical conditions that may occur during disease, such as a pH reduction or a change in the bile salt content (35), and are extremely sensitive to oxygen, although they can persist in environments with oxygen due to a flavin-thiol-dependent extracellular electron shuttle (61). These observations should be taken into account in the development of treatment strategies aiming to restore *F. prausnitzii* population in patients suffering from intestinal disorders. Novel treatments like prebiotics could be a strategy in order to boost the remaining *F. prausnitzii* populations in patients with gut disorders. The use of treatment strategies based on probiotic *F. prausnitzii* strains that are more tolerant of the gut conditions that prevail during intestinal disturbances could also be a suitable approach.

The observed imbalance in the prevalence of phylogroups between subjects provides an additional way to understand the role of this species in IBD since differences in their prevalence between healthy subjects and IBD patients must have biological relevance; eg., they may reflect differential responses to the host environment. Currently, no phenotypic trait consistently distinguishes *F. prausnitzii* members from one or the other phylogroup (35), but the existence of differences in members of different phylogroups due to, for instance, horizontal gene transfer, would not be surprising. Phenotypic differences between isolates from different phylogroups (which also have their characteristic 16S rRNA sequence) with respect to their capability to use carbohydrates from diet and/or host-derived, as well as their tolerance to bile salts and pH, have been observed (35). It can be hypothesized that differences in phylogroup composition reflect variations in sensitivity to such environmental factors or to interactions with the host, and it has been demonstrated that *F. prausnitzii* ATCC2768 (phylogroup I) and *F. prausnitzii* A2-165 (phylogroup II) are linked to the modulation of different urinary metabolites related to different host pathways (40). Our study does not allow us to decipher the biological relevance of the changes in population composition that were observed, but it points out that the *F. prausnitzii* populations hosted by different groups of subjects are different, and to address this question, further studies based on the isolation and characterization of *F. prausnitzii* isolates from subjects with these disorders and controls would be interesting.

Conclusion

Although members of the *F. prausnitzii* population are shared between healthy subjects and those with gut diseases, there is a loss of richness and a different distribution of specific phylotypes in IBD patients. The imbalance in phylogroups (OTU97_1 and OTU97_2) and the abundance of specific phylotypes can be used as biomarkers to distinguish some intestinal diseases, such as IBD or CRC.

Acknowledgements

This work was partially funded by the Spanish Ministry of Education and Science through project SAF2010-15896. Mireia Lopez-Siles was recipient of an FI grant from the Generalitat de Catalunya (2010FI_B2 00135), which receives support from the European Union Commissionate. Prof. Harry J. Flint and Dr. Sylvia H. Duncan acknowledge support from the Food, Land and People program of the Scottish government.

We thank Teresa Mas-de-Xaxars for assistance with recruitment of samples from CRC patients, Natàlia Adell from the Serveis Tècnics de Recerca for statistical assistance, Pau Boher for assistance with heat map analysis, and Marc Llirós and Carla Camprubí-Font for critically revising the manuscript. We appreciate the generosity of the patients who freely gave their time and samples to make this study possible, and the surgical theatre staff of all centers for their dedication and careful sample collection.

References

1. **Eckburg PB, Bik EM, Bernstein CN, Purdom E, Dethlefsen L, Sargent M, Gill SR, Nelson KE, Relman DA.** 2005. Diversity of the human intestinal microbial flora. Science **308:**1635-1638.
2. **Sekirov I, Russell SL, Antunes LC, Finlay BB.** 2010. Gut microbiota in health and disease. Physiol Rev **90:**859-904.
3. **Qin J, Li R, Raes J, Arumugam M, Burgdorf KS, Manichanh C, Nielsen T, Pons N, Levenez F, Yamada T, Mende DR, Li J, Xu J, Li S, Li D, Cao J, Wang B, Liang H, Zheng H, Xie Y, Tap J, Lepage P, Bertalan M, Batto JM, Hansen T, Le Paslier D, Linneberg A, Nielsen HB, Pelletier E, Renault P, Sicheritz-Ponten T, Turner K, Zhu H, Yu C, Jian M, Zhou Y, Li Y, Zhang X, Qin N, Yang H, Wang J, Brunak S, Dore J, Guarner F, Kristiansen K, Pedersen O, Parkhill J, Weissenbach J, Bork P, Ehrlich SD.** 2010. A human gut microbial gene catalogue established by metagenomic sequencing. Nature **464:**59-65.
4. **Nava GM, Stappenbeck TS.** 2011. Diversity of the autochthonous colonic microbiota. Gut Microbes **2:**99-104.
5. **Baumgart M, Dogan B, Rishniw M, Weitzman G, Bosworth B, Yantiss R, Orsi RH, Wiedmann M, McDonough P, Kim SG, Berg D, Schukken Y, Scherl E, Simpson KW.** 2007. Culture independent analysis of ileal mucosa reveals a selective increase in invasive *Escherichia coli* of novel phylogeny relative to depletion of Clostridiales in Crohn's disease involving the ileum. ISME J **1:**403-418.
6. **Arumugam M, Raes J, Pelletier E, Le Paslier D, Yamada T, Mende DR, Fernandes GR, Tap J, Bruls T, Batto JM, Bertalan M, Borruel N, Casellas F, Fernandez L, Gautier L, Hansen T, Hattori M, Hayashi T, Kleerebezem M, Kurokawa K, Leclerc M, Levenez F, Manichanh C, Nielsen HB, Nielsen T, Pons N, Poulain J, Qin J, Sicheritz-Ponten T, Tims S, Torrents D, Ugarte E, Zoetendal EG, Wang J, Guarner F, Pedersen O, de Vos WM, Brunak S, Dore J, Antolin M, Artiguenave F, Blottiere HM, Almeida M, Brechot C, Cara C, Chervaux C,**

Cultrone A, Delorme C, Denariaz G, Dervyn R, et al. 2011. Enterotypes of the human gut microbiome. Nature **473:**174-180.

7. **Hold GL, Schwiertz A, Aminov RI, Blaut M, Flint HJ.** 2003. Oligonucleotide probes that detect quantitatively significant groups of butyrate-producing bacteria in human feces. Appl Environ Microbiol **69:**4320-4324.
8. **Schwiertz A, Jacobi M, Frick JS, Richter M, Rusch K, Kohler H.** 2010. Microbiota in pediatric inflammatory bowel disease. J Pediatr **157:**240-244
9. **Suau A, Rochet V, Sghir A, Gramet G, Brewaeys S, Sutren M, Rigottier-Gois L, Doré J.** 2001. *Fusobacterium prausnitzii* and related species represent a dominant group within the human fecal flora. Systematic and Applied Microbiology **24:**139-145.
10. **Walker AW, Ince J, Duncan SH, Webster LM, Holtrop G, Ze X, Brown D, Stares MD, Scott P, Bergerat A, Louis P, McIntosh F, Johnstone AM, Lobley GE, Parkhill J, Flint HJ.** 2011. Dominant and diet-responsive groups of bacteria within the human colonic microbiota. ISME J**:**220-230.
11. **Swidsinski A, Loening-Baucke V, Lochs H, Hale LP.** 2005. Spatial organization of bacterial flora in normal and inflamed intestine: a fluorescence in situ hybridization study in mice. World J Gastroenterol **11:**1131-1140.
12. **Miquel S, Martin R, Rossi O, Bermudez-Humaran L, Chatel J, Sokol H, Thomas M, Wells J, Langella P.** 2013. *Faecalibacterium prausnitzii* and human intestinal health. Curr Opin Microbiol **16:**255-261.
13. **Frank DN, St Amand AL, Feldman RA, Boedeker EC, Harpaz N, Pace NR.** 2007. Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases. Proc Natl Acad Sci U S A **104:**13780-13785.
14. **Lopez-Siles M, Martinez-Medina M, Busquets D, Sabat-Mir M, Duncan SH, Flint HJ, Aldeguer X, Garcia-Gil LJ.** 2014. Mucosa-associated *Faecalibacterium prausnitzii* and *Escherichia coli* co-abundance can distinguish Irritable Bowel Syndrome and Inflammatory Bowel Disease phenotypes. Int J Med Microbiol **304:**464-475.
15. **Martinez-Medina M, Aldeguer X, Gonzalez-Huix F, Acero D, Garcia-Gil LJ.** 2006. Abnormal microbiota composition in the ileocolonic mucosa of Crohn's disease patients as revealed by polymerase chain reaction-denaturing gradient gel electrophoresis. Inflamm Bowel Dis **12:**1136-1145.
16. **Sokol H, Pigneur B, Watterlot L, Lakhdari O, Bermudez-Humaran LG, Gratadoux JJ, Blugeon S, Bridonneau C, Furet JP, Corthier G, Grangette C, Vasquez N, Pochart P, Trugnan G, Thomas G, Blottiere HM, Dore J, Marteau P, Seksik P, Langella P.** 2008. *Faecalibacterium prausnitzii* is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. Proc Natl Acad Sci U S A **105:**16731-16736.
17. **Sokol H, Seksik P, Furet JP, Firmesse O, Nion-Larmurier I, Beaugerie L, Cosnes J, Corthier G, Marteau P, Dore J.** 2009. Low counts of *Faecalibacterium prausnitzii* in colitis microbiota. Inflamm Bowel Dis **15:**1183-1189.
18. **Swidsinski A, Loening-Baucke V, Vaneechoutte M, Doerffel Y.** 2008. Active Crohn's disease and ulcerative colitis can be specifically diagnosed and monitored based on the biostructure of the fecal flora. Inflamm Bowel Dis **14:**147-161.
19. **Willing B, Halfvarson J, Dicksved J, Rosenquist M, Jarnerot G, Engstrand L, Tysk C, Jansson JK.** 2009. Twin studies reveal specific imbalances in the mucosa-associated microbiota of patients with ileal Crohn's disease. Inflamm Bowel Dis **15:**653-660.
20. **Furet JP, Kong LC, Tap J, Poitou C, Basdevant A, Bouillot JL, Mariat D, Corthier G, Dore J, Henegar C, Rizkalla S, Clement K.** 2010. Differential adaptation of human gut microbiota to bariatric surgery-induced weight loss: links with metabolic and low-grade inflammation markers. Diabetes **59:**3049-3057.
21. **Hansen R, Russell RK, Reiff C, Louis P, McIntosh F, Berry SH, Mukhopadhya I, Bisset WM, Barclay AR, Bishop J, Flynn DM, McGrogan P, Loganathan S, Mahdi G, Flint HJ, El-Omar EM, Hold GL.** 2012. Microbiota of de-novo pediatric IBD: increased *Faecalibacterium prausnitzii* and reduced bacterial diversity in Crohn's but not in ulcerative colitis. Am J Gastroenterol **107:**1913-1922.
22. **Jia W, Whitehead RN, Griffiths L, Dawson C, Waring RH, Ramsden DB, Hunter JO, Cole JA.** 2010. Is the abundance of *Faecalibacterium prausnitzii* relevant to Crohn's disease? FEMS Microbiol Lett **310:**138-144.
23. **Kabeerdoss J, Sankaran V, Pugazhendhi S, Ramakrishna BS.** 2013. *Clostridium leptum* group bacteria abundance and diversity in the fecal microbiota of patients with inflammatory bowel disease: a case-control study in India. BMC Gastroenterol **13:**20.
24. **Machiels K, Joossens M, Sabino J, De Preter V, Arijs I, Eeckhaut V, Ballet V, Claes K, Van Immerseel F, Verbeke K, Ferrante M, Verhaegen J, Rutgeerts P, Vermeire S.** 2013. A decrease of the butyrate-producing species *Roseburia hominis* and *Faecalibacterium prausnitzii* defines dysbiosis in patients with ulcerative colitis. Gut.

25. **McLaughlin SD, Clark SK, Tekkis PP, Nicholls RJ, Ciclitira PJ.** 2010. The bacterial pathogenesis and treatment of pouchitis. Therap Adv Gastroenterol **3:**335-348.
26. **Vermeiren J, Van den Abbeele P, Laukens D, Vigsnaes LK, De Vos M, Boon N, Van de Wiele T.** 2012. Decreased colonization of fecal *Clostridium coccoides/Eubacterium rectale* species from ulcerative colitis patients in an in vitro dynamic gut model with mucin environment. FEMS Microbiol Ecol **79:**685-696.
27. **Rajilic-Stojanovic M, Biagi E, Heilig HG, Kajander K, Kekkonen RA, Tims S, de Vos WM.** 2011. Global and deep molecular analysis of microbiota signatures in fecal samples from patients with irritable bowel syndrome. Gastroenterology **141:**1792-1801.
28. **Balamurugan R, Rajendiran E, George S, Samuel GV, Ramakrishna BS.** 2008. Real-time polymerase chain reaction quantification of specific butyrate-producing bacteria, *Desulfovibrio* and *Enterococcus faecalis* in the feces of patients with colorectal cancer. J Gastroenterol Hepatol **23:**1298-1303.
29. **Sobhani I, Tap J, Roudot-Thoraval F, Roperch JP, Letulle S, Langella P, Corthier G, Tran Van Nhieu J, Furet JP.** 2011. Microbial dysbiosis in colorectal cancer (CRC) patients. PLoS One **6:**e16393.
30. **de Goffau MC, Luopajarvi K, Knip M, Ilonen J, Ruohtula T, Harkonen T, Orivuori L, Hakala S, Welling GW, Harmsen HJ, Vaarala O.** 2013. Fecal microbiota composition differs between children with beta-cell autoimmunity and those without. Diabetes **62:**1238-1244.
31. **Qin J, Li Y, Cai Z, Li S, Zhu J, Zhang F, Liang S, Zhang W, Guan Y, Shen D, Peng Y, Zhang D, Jie Z, Wu W, Qin Y, Xue W, Li J, Han L, Lu D, Wu P, Dai Y, Sun X, Li Z, Tang A, Zhong S, Li X, Chen W, Xu R, Wang M, Feng Q, Gong M, Yu J, Zhang Y, Zhang M, Hansen T, Sanchez G, Raes J, Falony G, Okuda S, Almeida M, LeChatelier E, Renault P, Pons N, Batto JM, Zhang Z, Chen H, Yang R, Zheng W, Yang H, Wang J, et al.** 2012. A metagenome-wide association study of gut microbiota in type 2 diabetes. Nature **490:**55-60.
32. **Karlsson FH, Tremaroli V, Nookaew I, Bergstrom G, Behre CJ, Fagerberg B, Nielsen J, Backhed F.** 2013. Gut metagenome in European women with normal, impaired and diabetic glucose control. Nature **498:**99-103.
33. **Barcenilla A, Pryde SE, Martin JC, Duncan SH, Stewart CS, Henderson C, Flint HJ.** 2000. Phylogenetic relationships of butyrate-producing bacteria from the human gut. Appl Environ Microbiol **66:**1654-1661.
34. **Duncan SH, Hold GL, Harmsen HJ, Stewart CS, Flint HJ.** 2002. Growth requirements and fermentation products of *Fusobacterium prausnitzii*, and a proposal to reclassify it as *Faecalibacterium prausnitzii* gen. nov., comb. nov. Int J Syst Evol Microbiol **52:**2141-2146.
35. **Lopez-Siles M, Khan TM, Duncan SH, Harmsen HJ, Garcia-Gil LJ, Flint HJ.** 2012. Cultured representatives of two major phylogroups of human colonic *Faecalibacterium prausnitzii* can utilize pectin, uronic acids, and host-derived substrates for growth. Appl Environ Microbiol **78:**420-428.
36. **Louis P, Flint HJ.** 2009. Diversity, metabolism and microbial ecology of butyrate-producing bacteria from the human large intestine. FEMS Microbiol Lett **294:**1-8.
37. **Martin R, Chain F, Miquel S, Lu J, Gratadoux JJ, Sokol H, Verdu EF, Bercik P, Bermudez-Humaran LG, Langella P.** 2014. The commensal bacterium *Faecalibacterium prausnitzii* is protective in DNBS-induced chronic moderate and severe colitis models. Inflamm Bowel Dis **20:**417-430.
38. **Carlsson AH, Yakymenko O, Olivier I, Hakansson F, Postma E, Keita AV, Soderholm JD.** 2013. *Faecalibacterium prausnitzii* supernatant improves intestinal barrier function in mice DSS colitis. Scand J Gastroenterol **48:**1136-1144.
39. **Wrzosek L, Miquel S, Noordine ML, Bouet S, Chevalier-Curt MJ, Robert V, Philippe C, Bridonneau C, Cherbuy C, Robbe-Masselot C, Langella P, Thomas M.** 2013. *Bacteroides thetaiotaomicron* and *Faecalibacterium prausnitzii* influence the production of mucus glycans and the development of goblet cells in the colonic epithelium of a gnotobiotic model rodent. BMC Biol **11:**61.
40. **Li M, Wang B, Zhang M, Rantalainen M, Wang S, Zhou H, Zhang Y, Shen J, Pang X, Wei H, Chen Y, Lu H, Zuo J, Su M, Qiu Y, Jia W, Xiao C, Smith LM, Yang S, Holmes E, Tang H, Zhao G, Nicholson JK, Li L, Zhao L.** 2008. Symbiotic gut microbes modulate human metabolic phenotypes. Proc Natl Acad Sci U S A **105:**2117-2122.
41. **Tap J, Mondot S, Levenez F, Pelletier E, Caron C, Furet JP, Ugarte E, Munoz-Tamayo R, Paslier DL, Nalin R, Dore J, Leclerc M.** 2009. Towards the human intestinal microbiota phylogenetic core. Environ Microbiol **11:**2574-2584.
42. **Silverberg MS, Satsangi J, Ahmad T, Arnott ID, Bernstein CN, Brant SR, Caprilli R, Colombel JF, Gasche C, Geboes K, Jewell DP, Karban A, Loftus Jr EV, Pena AS, Riddell RH, Sachar DB, Schreiber S, Steinhart AH, Targan SR, Vermeire S, Warren BF.** 2005. Toward an integrated clinical, molecular and serological classification of inflammatory bowel disease: Report of a Working Party of the 2005 Montreal World Congress of Gastroenterology. Can J Gastroenterol **19 Suppl A:**5-36.

43. **Zoetendal EG, von Wright A, Vilpponen-Salmela T, Ben-Amor K, Akkermans AD, de Vos WM.** 2002. Mucosa-associated bacteria in the human gastrointestinal tract are uniformly distributed along the colon and differ from the community recovered from feces. Appl Environ Microbiol **68:**3401-3407.
44. **Thompson JD, Higgins DG, Gibson TJ.** 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res **22:**4673-4680.
45. **Maidak BL, Cole JR, Lilburn TG, Parker CT, Jr., Saxman PR, Farris RJ, Garrity GM, Olsen GJ, Schmidt TM, Tiedje JM.** 2001. The RDP-II (Ribosomal Database Project). Nucleic Acids Res **29:**173-174.
46. **Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res **25:**3389-3402.
47. **Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley GL, Vandesompele J, Wittwer CT.** 2009. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. Clin Chem **55:**611-622.
48. **Muyzer G, de Waal EC, Uitterlinden AG.** 1993. Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. Appl Environ Microbiol **59:**695-700.
49. **Hall TA.** 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucl Acids Symp Ser **41:**95-98.
50. **Wright ES, Yilmaz LS, Noguera DR.** 2012. DECIPHER, a search-based approach to chimera identification for 16S rRNA sequences. Appl Environ Microbiol **78:**717-725.
51. **Saitou N, Nei M.** 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol **4:**406-425.
52. **Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ, Sahl JW, Stres B, Thallinger GG, Van Horn DJ, Weber CF.** 2009. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. Appl Environ Microbiol **75:**7537-7541.
53. **Best WR, Becktel JM, Singleton JW, Kern F, Jr.** 1976. Development of a Crohn's disease activity index. National Cooperative Crohn's Disease Study. Gastroenterology **70:**439-444.
54. **Heuer H SK.** 1997. Application of denaturing gradient gel electrophoresis (DGGE) and temperature gradient gel electrophoresis (TGGE) for studying soil microbial communities. *In* van Elsas JD WE, Trevors JT (ed), Modern Soil Microbiology. NY:Marcel Dekker Inc, New York.
55. **Vetrovsky T, Baldrian P.** 2013. The variability of the 16S rRNA gene in bacterial genomes and its consequences for bacterial community analyses. PLoS One **8:**e57923.
56. **Stoddard SF, Smith BJ, Hein R, Roller BRK, Schmidt TM.** 2014. rrnDB: improved tools for interpreting rRNA gene abundance in bacteria and archaea and a new foundation for future development. Nucleic Acids Research.
57. **Nugent SG, Kumar D, Rampton DS, Evans DF.** 2001. Intestinal luminal pH in inflammatory bowel disease: possible determinants and implications for therapy with aminosalicylates and other drugs. Gut **48:**571-577.
58. **Pereira SP, Bain IM, Kumar D, Dowling RH.** 2003. Bile composition in inflammatory bowel disease: ileal disease and colectomy, but not colitis, induce lithogenic bile. Aliment Pharmacol Ther **17:**923-933.
59. **Morgan XC, Tickle TL, Sokol H, Gevers D, Devaney KL, Ward DV, Reyes JA, Shah SA, LeLeiko N, Snapper SB, Bousvaros A, Korzenik J, Sands BE, Xavier RJ, Huttenhower C.** 2012. Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment. Genome Biol **13:**R79.
60. **Seril DN, Liao J, Yang GY, Yang CS.** 2003. Oxidative stress and ulcerative colitis-associated carcinogenesis: studies in humans and animal models. Carcinogenesis **24:**353-362.
61. **Khan MT, Duncan SH, Stams AJ, van Dijl JM, Flint HJ, Harmsen HJ.** 2012. The gut anaerobe *Faecalibacterium prausnitzii* uses an extracellular electron shuttle to grow at oxic-anoxic interphases. ISME J **6:**1578-1585.

Chapter 3‡

Abundance of mucosa-associated *Faecalibacterium prausnitzii* and usefulness as diagnostic biomarker in inflammatory bowel diseases

Significance of this study

What was already known on this subject?

- Mounting evidences suggest that *F. prausnitzii* is reduced in several intestinal disorders, especially in some CD subtypes. Two phylogroups exist within *F. prausnitzii*, and their prevalence differs between H subjects and patients with IBD and CRC.
- Other species such as *E. coli* have been consistently reported as dysbiosis-signature of CD patients.

What are the new findings?

- Total *F. prausnitzii* and phylogroup I are depleted in CD, UC and CRC patients in comparison to H subjects, whilst phylogroup II is specifically reduced in CD. Within IBD, those CD patients with ileal involvement have the lowest *F. prausnitzii* abundances, whereas those with colonic CD have values similar to UC patients.
- *F. prausnitzii* abundance does not change by medication, and remains reduced in IBD patients regardless of their activity status at the moment of colonoscopy. There are no differences in *F. prausnitzii* load between patients with and without intestinal resection.

How might this impact on research in the foreseeable future?

- Quantification of *F. prausnitzii* and *E. coli* may help to identify gut disorders, and to classify inflammatory bowel disease locations.

---

‡

Part of this chapter has been published in the *International Journal of Medical Microbiologylogy*: Lopez-Siles, M *et al.* **Mucosa-associated *Faecalibacterium prausnitzii* and *Escherichia coli* co-abundance can distinguish Irritable Bowel Syndrome and Inflammatory Bowel Disease phenotypes.** *Int J Med Microbiol* 2014; 304:464-475

Part of this chapter has been saccepted in the journal *Inflammatory Bowel Diseases on 3rd August 2015*. Lopez-Siles, M *et al.* **Changes in the abundance of *Faecalibacterium prausnitzii* phylogroups I and II in the intestinal mucosa of inflammatory bowel disease and colorectal cancer patients.** doi:10.1097/MIB.0000000000000590

Article III

International Journal of Medical Microbiology 304 (2014) 464–475

Contents lists available at ScienceDirect

International Journal of Medical Microbiology

journal homepage: www.elsevier.com/locate/ijmm

Mucosa-associated *Faecalibacterium prausnitzii* and *Escherichia coli* co-abundance can distinguish Irritable Bowel Syndrome and Inflammatory Bowel Disease phenotypes

Mireia Lopez-Siles[a], Margarita Martinez-Medina[a], David Busquets[b], Miriam Sabat-Mir[c], Sylvia H. Duncan[d], Harry J. Flint[d], Xavier Aldeguer[b], L. Jesús Garcia-Gil[a,*]

[a] *Laboratory of Molecular Microbiology, Biology Department, Universitat de Girona, Girona, Spain*
[b] *Departament de Gastroenterologia, Hospital Dr. Josep Trueta, Girona, Spain*
[c] *Departament de Gastroenterologia, Hospital Santa Caterina, Salt, Girona, Spain*
[d] *Microbiology Group, Rowett Institute of Nutrition and Health, University of Aberdeen, Aberdeen, United Kingdom*

ARTICLE INFO

Article history:
Received 22 July 2013
Received in revised form 31 October 2013
Accepted 9 February 2014

Keywords:
Faecalibacterium prausnitzii
Escherichia coli
Inflammatory Bowel Disease
Irritable Bowel Syndrome
Diagnostics
Prognostics

ABSTRACT

Background: Crohn's disease (CD) and ulcerative colitis (UC) diagnosis requires comprehensive examination of the patient. *Faecalibacterium prausnitzii* and *Escherichia coli* have been reported as representatives of Inflammatory Bowel Disease (IBD) dysbiosis. The aim was to determine whether or not quantification of these species can be used as a complementary tool either for diagnostic or prognostic purposes.
Methods: Mucosa-associated *F. prausnitzii* and *E. coli* abundance was determined in 28 controls (H), 45 CD, 28 UC patients and 10 irritable bowel syndrome (IBS) subjects by quantitative polymerase chain reaction (qPCR) and the *F. prausnitzii-E. coli* index (F-E index) was calculated. Species abundances were normalized to total bacteria and human cells. Data was analyzed taking into account patients' phenotype and most relevant clinical characteristics.
Results: IBD patients had lower *F. prausnitzii* abundance than H and IBS ($P < 0.001$). CD patients showed higher *E. coli* counts than H and UC patients ($P < 0.001$). The F-E index discriminated between H, CD and UC patients, and even between disease phenotypes that are usually difficult to distinguish as ileal-CD (I-CD) from ileocolonic-CD and colonic-CD from extensive colitis. *E. coli* increased in active CD patients, and remission in I-CD patients was compromised by high abundance of this species. Treatment with anti-tumor necrosis factor (TNF) α diminished *E. coli* abundance in I-CD whereas none of the treatments counterbalanced *F. prausnitzii* depletion.
Conclusion: *F. prausnitzii* and *E. coli* are useful indicators to assist in IBD phenotype classification. The abundance of these species could also be used as a supporting prognostic tool in I-CD patients. Our data indicates that current medication does not restore the levels of these two species to those found in a healthy gut.

Introduction

Inflammatory bowel disease (IBD) comprises a group of idiopathic, chronic, inflammatory intestinal disorders. Its two most important disease categories are Crohn's disease (CD) and ulcerative colitis (UC) (Baumgart and Carding, 2007; Baumgart and Sandborn, 2007; Xavier and Podolsky, 2007). Although both intestinal diseases differ in terms of their location, the distribution of inflamed areas and their histology, classification of these disease states can be difficult given their overlapping clinical and pathological characteristics (Yantiss and Odze, 2006). To clearly discriminate both diseases is essential to establish an appropriate treatment strategy. In addition, other digestive disorders such as irritable bowel syndrome (IBS) can mimic IBD clinically, particularly in the early stages, increasing its likelihood of misdiagnosis (Bernstein et al., 2010; Nikolaus and Schreiber, 2007).

Given the absence of pathognomonic features, the diagnosis for IBD currently requires a comprehensive examination of the patient that includes clinical, endoscopic, radiologic, and histological criteria (Bernstein and Shanahan, 2008). IBD is an intermittent disease,

* Corresponding author at: Laboratory of Molecular Microbiology, Biology Department, Universitat de Girona, Carrer de Maria Aurèlia Capmany, 69, E-17071 Girona, Spain. Tel.: +34 972 418175; fax: +34 972 418150.
E-mail address: jesus.garcia@udg.edu (L.J. Garcia-Gil).

http://dx.doi.org/10.1016/j.ijmm.2014.02.009
1438-4221/

whose clinical manifestations are hardly predictable and unstable during its course. Symptoms range from mild to severe during relapses and may disappear or decrease during episodes of remission. Thus, careful consideration of a patient's clinical data and a long monitoring period are necessary to accurately classify the disease phenotype (Bernstein et al., 2010; Louis et al., 2001).

Although the pathogenesis of IBD is incompletely understood, it is known that it is a complex disease in which many factors determine who develops IBD, the age of presentation, and the specific manifestations of disease (Bernstein et al., 2010; Kaser et al., 2010; Manichanh et al., 2012). Currently, the most generally accepted hypothesis is that genetic and environmental factors such as altered luminal bacteria and enhanced intestinal permeability play a role in the deregulation of intestinal immunity, which in turn may lead to gastrointestinal injury (Sartor, 2006; Xavier and Podolsky, 2007).

The role of the gut microbiota in the onset and perpetuation of intestinal inflammation in IBD has been a topic systematically studied during the last 10 years (for review see (De Cruz et al., 2012; Elson and Cong, 2012; Manichanh et al., 2012) and references therein). It is well established by studies performed both in fecal or mucosa-associated communities, either by culture-dependent or molecular methods that CD patients have an altered microbiota, which differs from that found in patients with UC and as well as of that in healthy controls (Manichanh et al., 2012). This dysbiosis is characteristic of the disease as it is not shared with unaffected monozygotic twins or relatives despite the common genetic background and the shared environment (Joossens et al., 2011; Willing et al., 2009). Although the reported changes are not always consistent, most studies agree that numbers of *Firmicutes*, particularly the species *Faecalibacterium prausnitzii*, are depleted in patients with CD (Frank et al., 2007; Martinez-Medina et al., 2006; Miquel et al., 2013; Sokol et al., 2009; Swidsinski et al., 2008; Willing et al., 2009) whereas *Proteobacteria*, especially *Escherichia coli*, are increased predominantly in CD patients with ileal involvement (Martinez-Medina et al., 2009; Mondot et al., 2011; Seksik et al., 2003; Willing et al., 2009). Taken together these findings indicate that the abundance of these two bacterial groups might be a reliable indicator of dysbiosis in CD patients.

Application of molecular methods to specifically monitor changes of key microorganisms in the gut is of particular interest, since it may provide an innovative source of additional information to assist clinicians in disease diagnosis and management. To our knowledge, few studies have been conducted with this aim in respect of IBD. Interestingly, a reduction in *F. prausnitzii* abundance has however been correlated with IBD patients' activity, flare ups and remission state (Sokol et al., 2009), but few studies have addressed the question of whether this bacterium or other key dysbiosis representatives could be useful to assist IBD diagnostics or to monitor disease progression. Swidsinski and colleagues have reported that CD and UC could be diagnosed through monitoring *F. prausnitzii* abundance in conjunction with fecal leucocyte counts (Swidsinski et al., 2008). Recently new phylogenetic specificities of CD microbiota have been highlighted by identifying a set of six species discriminatory for CD patients with ileal involvement, which also provides a preliminary diagnostic tool (Mondot et al., 2011). However, further analysis including all CD and UC phenotypes should be performed in order to determine the extent of dysbiosis within all disease categories. In addition, comprehensive studies are lacking to show how patients' clinical data correlates with changes in the abundance of these bacterial indicators, and how the different therapies may affect the abundance of these species.

This work aims at testing whether or not mucosa-associated *F. prausnitzii* and *E. coli* abundances could be used to differentially diagnose IBD patients and monitor the evolution of the disease. To achieve this objective, the abundance of both bacterial species was determined in CD, UC and IBS patients and in healthy controls. A novel multiplex quantitative polymerase chain reaction (qPCR) assay was developed for *F. prausnitzii*, valid for the quantification of the two known phylogroups within this species. Furthermore, data were analyzed taking into account patients' most relevant clinical characteristics, in order to determine its usefulness to predict disease progression. Medication at sampling was also considered in order to determine whether any of the current therapies are effective in correcting this dysbiosis.

Materials and methods

Patients, clinical data and sampling

A Spanish cohort consisting of 73 IBD patients, including 45 CD and 28 UC has been compared with those from ten IBS patients and 28 healthy control subjects (H). Subjects were recruited by the Gastroenterology Services of the Hospital Universitari Dr. Josep Trueta (Girona, Spain) and the Hospital Santa Caterina (Salt, Spain). Patients were sex- and age-matched, except CD patients who were significantly younger than those in the H and IBS groups ($P < 0.001$) (Table 1). IBD patients were diagnosed according to standard clinical, pathological and endoscopic criteria, were categorized according to the Montreal classification (Silverberg et al., 2005), and clinically relevant data was collected. IBS patients were diagnosed according to Rome III criteria (available at http://www.romecriteria.org/criteria/). The control group consisted of subjects with normal colonoscopy who underwent this procedure for different reasons as rectorrhagia ($N = 9$), colorectal cancer familial history ($N = 10$), and abdominal pain ($N = 9$). None of the subjects received antimicrobial treatment for at least two months before colonoscopy.

Prior to colonoscopy, patients were subjected to cleansing of the gastrointestinal tract using Casenglicol® following manufacturer's guidelines. During routine endoscopy, up to three biopsy samples per patient were taken from different locations along the gut (distal ileum, colon, and rectum) following standard procedures. For IBD patients, additional samples from ulcerated and non ulcerated mucosa according to macroscopic criteria were taken when technically possible. All biopsies were immediately placed in sterile tubes without any buffer and stored at −20 °C following completion of the whole endoscopic procedure, for each patient. DNA extraction was then performed on these samples within the following 6 months.

A subgroup of 10 CD patients who started adalimumab therapy (HUMIRA; Abbott Laboratories, Chicago, IL) was enrolled on a follow-up study and rectal samples were also collected one and three months after the first colonoscopy.

Ethical considerations

This work was approved by the Ethics Committee of Clinical Research of the Hospital Universitari Dr. Josep Trueta (Girona, Spain) and the Institut d'Assistència Sanitària of Girona (Salt, Spain) on 24th February 2009 and 21st April 2009, respectively. Informed consent from the subjects was obtained before enrollment.

Sample treatment and DNA extraction

Prior to DNA extraction, biopsies were subjected to two mild ultrasound wash cycles to discard transient and loosely attached bacteria as previously reported (Martinez-Medina et al., 2006). DNA was extracted using the NucleoSpin® Tissue Kit (Macherey-Nagel GmbH &Co., Germany). The support protocol for Gram positive bacteria and the RNAse treatment step were carried out. Genomic DNA was stored at −80 °C until use.

Table 1
Sample size and clinical characteristics of subjects.

	Healthy controls*	IBD		Irritable bowel syndrome (IBS)
		Crohn's disease	Ulcerative colitis	
N (patients)	28	45	28	10
Age (mean years ± SD)	47.1 ± 16.0	34.4 ± 11.2	40.5 ± 15.2	43.8 ± 10.8
Male (N, %)	16 (57.1%)	23 (50.0%)	17 (60.7%)	2 (20.0%)
Active (N, %)	na	28 (60.9%)	21 (75.0%)	nd
Previous surgery (N, %)	0	10 (21.7%)	1 (3.6%)	nd
Smokers (N, %)	nd	12 (26.1%)	2 (7.1%)	0
Treatment (N, %)**				
No treatment or mesalazine	na	17 (37.0%)	17 (60.7%)	nd
Moderate immunosuppressant	na	17 (37.0%)	4 (14.3%)	nd
Anti-TNFα (infliximab, adalimumab)	na	11 (23.9%)	5 (17.9%)	nd
CD Montreal classification				
*Age of diagnosis (N, %)***				
diag < 16y (A1)	na	5 (10.9%)	1 (3.6%)	na
diag 17–40y (A2)	na	33 (71.7%)	12 (42.9%)	na
diag >41y (A3)	na	6 (13.0%)	11 (39.3%)	na
Location (N, %)				
Ileal-CD (L1)	na	19 (41.3%)	na	na
Colonic-CD (L2)	na	13 (28.3%)	na	na
Ileocolonic-CD (L3)	na	13 (28.3%)	na	na
Behavior (N, %)**				
Non-stricturing, non-penetrating (B1)	na	31 (67.4%)	na	na
Stricturing (B2)	na	7 (15.2%)	na	na
UC classification (N, %)**				
Ulcerative proctitis (E1)	na	na	6 (21.4%)	na
Distal UC (E2)	na	na	13 (46.4%)	na
Extensive UC or pancolitis (E3)	na	na	7 (25.0%)	na
IBS subtype (N, %)**				
Diarrhea predominant type	na	na	na	2 (20.0%)
Constipation predominant type	na	na	na	2 (20.0%)

IBD, inflammatory bowel disease; IBS, Irritable bowel syndrome; TNF, tumour necrosis factor; nd, not determined; na, not applicable.
* Controls consisted of subjects who underwent colonoscopy for different reasons: 9/28 rectorrhagia, 10/28 colorectal cancer familial history and 9/28 abdominal pain.
** Medical treatment at the time of sampling was available in 26/28 UC patients; Age of disease onset was available for 44/45 CD patients, and 24/28 UC patients; Disease behavior at last follow-up before the time of sampling was available in 38/45 CD patients, and none had penetrating CD (B3); Maximal disease extent at the time of sampling was available in 26/28 UC patients; disease subtype was available in 4/10 Irritable bowel syndrome patients, and none had alternating predominant type.

DNA concentration and optical density ratios at 260/280 nm and 230/260 nm to check the purity of the extracts were determined with a NanoDrop ND-100 spectrophotometer (NanoDrop Technologies, USA).

Bacterial strains, growth conditions and DNA extraction from pure cultures

F. prausnitzii strains were from stocks held by the authors (Rowett Institute of Nutrition and Health, Aberdeen, United Kingdom) and several came from previous studies (Barcenilla et al., 2000; Cato et al., 1974; Duncan et al., 2002; Lopez-Siles et al., 2012; Louis et al., 2004). Additional bacterial strains were either available in our laboratory collection or were otherwise obtained from several biological resource centers specified in Table S2. When possible, bacteria were cultured aerobically or anaerobically on the recommended medium. DNA was extracted and purified by using the Wizard™ Genomic Purification Kit (Promega Corporation, USA) following the manufacturer's guidelines.

Quantification standards for quantitative PCR

Quantification standards of the *F. prausnitzii* DSM 17677 and *E. coli* CECT 105 16S rRNA genes were prepared in a genetic construct. The whole 16S rRNA gene of the target species were amplified by conventional PCR as previously reported (Lane, 1991; Weisburg et al., 1991) and further introduced in a pCR®4-TOPO® cloning plasmid by using the TOPO TA Cloning® Kit for sequencing (Invitrogen, CA, USA) following the manufacturer's guidelines. Plasmids were extracted using the NucleoSpin® Plasmid (Macherey-Nagel GmbH&Co., Germany). Inserts were further confirmed by sequencing using the Big Dye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, USA) on an ABI Prism 3130 automated DNA sequencer (Applied Biosystems, Foster City, CA, USA). Purified plasmids were linearized with *SpeI* (*F. prausnitzii*) or *PstI* (*E. coli*), and DNA quantified as detailed above. Initial target concentration was inferred considering the theoretical molecular weight (3.58×10^6 Da) and size (5421 bp) of the construct. Standard curves were obtained from 10-fold serial dilutions of the titrated suspension of linearized plasmids, and ranged from 100 to 10^7 copies/reaction, which correspond to the linear range span for all the reactions. As it is recommended to use the same standard for species-specific and group-specific primers and probe sets (Suzuki et al., 2000), the standard curve built for *F. prausnitzii* quantification was used for the total bacterial 16S rRNA gene quantification. Total bacteria 16S rRNA gene quantification and the *F. prausnitzii* standard curve were used to check the *E. coli* standard curve quantification in order to make sure that results obtained with both standard curves were comparable. For human cells, ten-fold serial dilutions of the human Xsomal DNA (Eurogentec, Belgium) were used to obtain the standard curve.

Quantitative PCR conditions

The species-specific 16S rRNA gene-targeted primers and probes used in this study are shown in Table 2. The abundance of *F. prausnitzii* was determined by using a novel assay, designed following the guidelines set by Applied Biosystems (Foster City, CA, USA) for the design of primers and probes, and taking into account the inclusion of both *F. prausnitzii* phylogroups (see details described in the supplemental material, according to the MIQE guidelines (Bustin et al., 2009)). The amplification reactions were carried out

Table 2
16S rRNA-targeted primers and probes used in this study.

Target	Primer and probe*	Sequence 5′–3′	Reference
Bacteria	F_Bact 1369	CGGTGAATACGTTCCCGG	Furet et al., 2009
	R_Prok_1492	TACGGCTACCTTGTTACGACTT	
	P_TM_1389F	**6FAM-CTTGTACACACCGCCCGTC-TAMRA**	
E. coli	E.coli 395 F	CATGCCGCGTGTATGAAGAA	Huijsdens et al., 2002
	E.coli 490 R	CGGGTAACGTCAATGAGCAAA	
	E.coli 437 PR	**6FAM-TATTAACTTTACTCCCTTCCTCCCCGCTGAA-TAMRA**	
F. prausnitzii	Fpra 428 F	TGTAAACTCCTGTTGTTGAGGAAGATAA	This study
	Fpra 583 R	GCGCTCCCTTTACACCCA	
	Fpra 493 PR	**6FAM-CAAGGAAGTGACGGCTAACTACGTGCCAG-TAMRA**	
DNA IAC‡	IAC F	TACGGATGAGGAGGACAAAGGA	This study
	IAC R	CACTTCgCTCTgATCCATTgg	
	IAC PR	**VIC®-CGCCGCTATGGGCATCGCA-TAMRA**	

* Probe sequences are in bold. P_TM1389F, E. coli 437 PR and Fpra493PR probes were 5′-labelled with FAM™ (6-carboxyfluorescin) as the reporter dye, whereas the IAC probe was 5′ labeled with VIC® (6-carboxyrhodamine) as reporter dye to allow multiplex detection. TAMRA™ was used as quencher dye at the 3′-end for all the probes.
‡ IAC, Internal Amplification Control; DNA IAC sequence (5′–3′): TACGGATGAGGAGGACAAAGGACGCCGCTATGGGCATCGCACCAATGGATCAGAGCGAAGTG.

in a total volume of 20 μl containing: 1× TaqMan® Universal PCR Master Mix 2× (Applied Biosystems, Foster City, CA, USA), 300 nM of each primer and 200 nM of each probe, 10^3 copies of an internal amplification control (IAC) template and up to 50 ng of genomic DNA template.

Previously reported 16S rDNA-targeting primers and probe were used for *E. coli* (Huijsdens et al., 2002) and total bacteria (Furet et al., 2009) quantifications, and amplification reactions were carried out as previously described (Martinez-Medina et al., 2009; Furet et al., 2009). Human cell numbers were determined with the control kit RT-CKFT-18S (Eurogentec, Belgium) according to manufacturer's instructions. All primers and hydrolysis probes were purchased from Applied Biosystems (Foster City, CA, USA). The IAC's DNA was synthesized by Bonsai technologies group (Alcobendas, Spain).

Samples were quantified in duplicate. For data analysis, the mean of the duplicate quantifications was used. Duplicates were considered valid if the standard deviation between quantification cycles (C_q) was <0.34 (i.e. a difference of <10% of the quantity was tolerated). Quantification controls to assess inter-run reproducibility were performed consisting of at least five reactions with a known number of target genes. Inhibition was tested by addition of an IAC in each reaction. It was considered that there was no inhibition if the obtained C_q was <0.34 different from those obtained when quantifying the IAC alone for any of the replicates. A no-template control consisting of a reaction without target (*F. prausnitzii, E. coli* or human) DNA template as well as a non-amplification control which did not contain any DNA template (either bacterial, human or IAC) were also included in each run. Negative controls resulted in undetectable C_q values in all cases.

All quantitative PCR were performed using a 7500 Real Time PCR system (Applied Biosystems, Foster City, CA, USA). The thermal profile was: a first step at 50 °C during 2 min for amperase treatment, followed by a 95 °C hold for 10 min to denature DNA and activate Ampli-Taq Gold polymerase, and a further 40 cycles consisting of a denaturation step at 95 °C for 15 seconds followed by an annealing and extension step at 60 °C for 1 min. Data was collected and analyzed with the 7500 SDS system software version 1.4 (Applied Biosystems, Foster City, CA, USA). The PCR efficiency ranged between 80 and 100% in all the reactions.

Sample size, data normalization, F. prausnitzii-E. coli index and statistical analysis

Sample size was defined taking into account the number of patients analysed in similar studies of bacterial abundance in patients suffering of these conditions (Frank et al., 2007; Martinez-Medina et al., 2006; Sokol et al., 2009; Swidsinski et al., 2008; Willing et al., 2009).

F. prausnitzii and *E. coli* 16S rRNA gene copy numbers were normalized to the total bacteria 16S rRNA gene. Data is given as $\log_{10}$ 16S rRNA gene copies of the target microorganism per million of bacterial 16S rRNA genes detected in the same sample. The *F. prausnitzii-E. coli* index (F-E index) was calculated as $[\log_{10}(F/Hc) - \log_{10}(E/Hc)/\log_{10}(TB/Hc)$, being F the 16S rRNA gene copies of *F. prausnitzii*, E the 16S rRNA gene copies of *E. coli*, Hc a million of human cells, and TB a million of 16S rRNA gene copies of total bacteria. This index allows the normalization of the biopsy size by quantifying human cells and includes total bacteria as an additional parameter, as it has been reported that it can vary between groups of patients (Kleessen et al., 2002; Schultsz et al., 1999; Swidsinski et al., 2002).

The variation coefficient was calculated as a measure of dispersion between samples from the same patient. As within a patient there were high differences between samples from different zones along the intestinal tract, analyses pooling all the biopsies together and separated by location were performed. The non-parametric Kruskal–Wallis test was used to test differences in variables with more than two categories (i.e. diagnostics, CD and UC phenotypes, and current medication). Pairwise comparisons of subcategories of these variables were further analyzed using a Mann–Whitney U test. This test was also used to compare, within a subgroup of patients variables with two categories as activity (active CD and UC patients when CDAI > 150 (Best et al., 1976) and a Mayo score >3, respectively), and intestinal resection.

Spearman correlation coefficient and significance between the two species quantities was calculated. The same statistics were used to analyze the correlation between each one of the species and the F-E index with respect to simple endoscopic score for CD (SES-CD), Mayo endoscopic score for UC (Pineton de Chambrun et al., 2010), C-reactive protein, and months to flare up in inactive IBD patients.

The receiver operating characteristic (ROC) curve analysis, a plot of the true positive rate (sensitivity) versus false positive rate (1-specificity), was applied to establish the usefulness of *F. prausnitzii*, *E. coli* and the F-E index to distinguish amongst different intestinal disorders. The accuracy of discrimination was measured by the area under the ROC curve (AUC). An AUC approaching 1 indicates that the test is highly sensitive as well as highly specific whereas an AUC approaching 0.5 indicates that the test is neither sensitive nor specific.

Table 3
Abundances of mucosa-associated *F. prausnitzii*, *E. coli* and F-E index in controls (H), Irritable Bowel Syndrome (IBS), Ulcerative Colitis (UC), and Crohn's disease (CD) patients. Disease phenotypes of UC and CD patients are analyzed as independent groups.

	n patients (n biopsies)	*F. prausnitzii*[*§]	*E. coli*[*§]	*F-E index*[*†]
H	**28 (59)**	**5.41 ± 0.55[a]**	**4.05 ± 1.18[a]**	**0.22 ± 0.21[a]**
IBS	**10 (26)**	**5.34 ± 0.57[a]**	**3.30 ± 1.13[a,b]**	**0.29 ± 0.17[a,b]**
UC	**28 (66)**	**4.95 ± 0.63[b]**	**3.04 ± 1.22[b]**	**0.30 ± 0.19[a]**
Location				
Ulcerative proctitis (E1)	6 (18)	5.12 ± 0.31[a,b]	3.04 ± 0.75[b]	0.33 ± 0.09[b]
Distal UC (E2)	13 (35)	4.44 ± 0.62[c]	2.92 ± 1.31[b]	0.33 ± 0.22[a,b]
Extensive UC or pancolitis (E3)	7 (13)	5.24 ± 0.68[a,b]	4.57 ± 1.43[a,b,c]	0.18 ± 0.15[a]
CD	**46 (91)**	**4.30 ± 1.28[c]**	**4.51 ± 1.08[c]**	**-0.02 ± 0.28[c]**
Location				
Ileal-CD (L1)	19 (39)	3.84 ± 1.38[d]	4.58 ± 1.11[c,d]	-0.19 ± 0.29[d]
Colonic-CD (L2)	13 (28)	5.08 ± 0.93[b,c]	4.58 ± 0.91[c]	0.01 ± 0.20[c]
Ileocolonic-CD (L3)	13 (24)	4.44 ± 1.01[c,d]	3.85 ± 1.23[a,b,c,d]	-0.02 ± 0.29[c]
Behavior[*]				
Non-stricturing, non-penetrating (B1)	31 (64)	4.26 ± 1.30[c]	4.35 ± 1.00[a,c]	0.00 ± 0.26[c]
Stricturing (B2)	7 (17)	3.52 ± 0.97[d]	5.25 ± 1.25[d]	-0.24 ± 0.25[d]

[*] Homogeneous subgroups (P > 0.05) within each variable (column) are indicated with the same superscript.
[§] Median $\log_{10}$ 16S rRNA gene copies/million bacterial 16S rRNA gene copies ± standard deviations.
[†] Median F-E index ± standard deviations. F-E index has been calculated as [(*F. prausnitzii* $\log_{10}$ 16S rRNA gene copies/million human cells)-(*E. coli* $\log_{10}$ 16S rRNA gene copies/million human cells)]/(total bacteria $\log_{10}$ 16S rRNA gene copies/million human cells).

Table 4
Abundances of mucosa-associated *F. prausnitzii*, *E. coli* and F-E index by zone of the gastrointestinal tract (ileum, colon and rectum) in controls (H), Irritable Bowel Syndrome (IBS), Ulcerative Colitis (UC), and Crohn's disease (CD) patients. Disease phenotypes of UC and CD patients are analyzed as independent groups.

	N biopsies	*F. prausnitzii*[*§]	*E. coli*[*§]	F-E index[*†]
Ileum				
H	**15**	**5.51 ± 0.53[a]**	**4.07 ± 1.23[a,b]**	**0.22 ± 0.15[a]**
IBS	**6**	**4.86 ± 1.43[ab]**	**4.92 ± 1.39[a,b,c]**	**-0.01 ± 0.16[b,c]**
UC				
Ulcerative proctitis (E1)	6	5.18 ± 0.23[a,b]	3.18 ± 0.87[a]	0.31 ± 0.09[a,b]
Distal UC (E2)	9	5.36 ± 0.54[a,b]	2.97 ± 1.15[a]	0.37 ± 0.09[a]
Extensive UC or pancolitis (E3)	5	5.24 ± 0.69[a,b]	4.89 ± 1.47[a,b,c]	0.14 ± 0.14[b,c]
CD				
Ileal-CD (L1)	11	3.96 ± 1.35[b]	4.96 ± 0.87[b,c]	-0.16 ± 0.25[c]
Colonic-CD (L2)	7	5.03 ± 1.00[a,b]	4.32 ± 0.85[a]	0.11 ± 0.23[a,b,c]
Ileocolonic-CD (L3)	5	3.33 ± 1.36[a,b]	5.26 ± 0.89[b,c]	-0.23 ± 0.26[c]
Colon				
H	**33**	**5.46 ± 0.63[a]**	**4.08 ± 1.24[a,d]**	**0.21 ± 0.24[a,b]**
IBS	**10**	**5.46 ± 0.19[a]**	**3.19 ± 1.31[a,b]**	**0.33 ± 0.16[a,b]**
UC				
Ulcerative proctitis (E1)	6	5.11 ± 0.17[b]	3.04 ± 0.77[b]	0.32 ± 0.10[a]
Distal UC (E2)	13	4.42 ± 0.61[b,c]	2.97 ± 1.51[b]	0.28 ± 0.25[a,b]
Extensive UC or pancolitis (E3)	7	5.12 ± 0.71[a,b,c]	3.45 ± 1.48[a,b,d]	0.20 ± 0.16[a,b]
CD				
Ileal-CD (L1)	19	2.74 ± 1.30[d]	4.55 ± 1.03[c,d]	-0.26 ± 0.29[d]
Colonic-CD (L2)	13	4.84 ± 0.85[c]	4.93 ± 0.68[c]	-0.01 ± 0.19[c]
Ileocolonic-CD (L3)	13	4.49 ± 1.07[a,b,c]	3.85 ± 1.30[d]	0.13 ± 0.32[c]
Rectum				
H	**11**	**5.28 ± 0.33[a]**	**3.86 ± 1.00[a]**	**0.22 ± 0.17[a,b]**
IBS	**10**	**5.31 ± 0.31[a]**	**3.32 ± 1.70[a,b]**	**0.28 ± 0.25[a,b]**
UC				
Ulcerative proctitis (E1)	6	5.13 ± 0.49[a]	3.19 ± 0.66[a,b]	0.33 ± 0.05[a]
Distal UC (E2)	13	4.49 ± 0.68[a,c]	2.55 ± 0.86[b]	0.33 ± 0.14[a]
Extensive UC or pancolitis (E3)	1	5.76[a,b,c]	4.76[a,b]	0.18[a,b,c]
CD				
Ileal-CD (L1)	9	4.25 ± 1.51[c]	4.01 ± 1.38[a,b]	0.01 ± 0.32[c]
Colonic-CD (L2)	8	5.09 ± 1.12[a,b,c]	4.53 ± 1.14[a]	0.04 ± 0.16[b,c]
Ileocolonic-CD (L3)	6	4.05 ± 0.32[b,c]	3.36 ± 0.50[a,b]	0.13 ± 0.10[c]

[*] Homogeneous subgroups (P > 0.05) within each variable (column) are indicated with the same superscript.
[§] Median $\log_{10}$ 16S rRNA gene copies/million bacterial 16S rRNA gene copies ± standard deviations.
[†] Median F-E index ± standard deviations. F-E index has been calculated as [(*F. prausnitzii* $\log_{10}$ 16S rRNA gene copies/million human cells)-(*E. coli* $\log_{10}$ 16S rRNA gene copies/million human cells)]/(total bacteria $\log_{10}$ 16S rRNA gene copies/million human cells).

All the statistical analyses were conducted via the SPSS 15.0 statistical package for Windows (LEAD Technologies, Inc.). Significance levels were established for P values ≤ 0.05.

Results

Features of the novel multiplex quantitative PCR assay for F. prausnitzii (both phylogroups)

In this study, a novel primer set and probe to quantify *F. prausnitzii* has been developed (Table 2, supplemental material), taking into account that it should equally detect and quantify the two recently described phylogroups of this species (Lopez-Siles et al., 2012). Additionally, an IAC has been included in order to report quantitative errors or false negative reactions due to inhibition, thus ensuring accurate quantification when using the assay for the analysis of clinical samples. The assay is totally specific, as assessed both in silico and in vitro with an average efficiency of 86%. The theoretical detection limit is of 106.6 16S rRNA genes of *F. prausnitzii* per reaction and allows quantification over a linear range span of at least 7 logarithms, starting at 10^3 target genes per reaction. The tool hereby developed is suitable to be applied for determinations of *F. prausnitzii* in human biopsy samples, considering that healthy persons harbor around 1.7×10^5 *F. prausnitzii*·mg tissue^{-1} (Ahmed et al., 2007).

Abundance of mucosa-associated F. prausnitzii and E. coli in healthy subjects, IBS and IBD patients by disease phenotype

The abundance of *F. prausnitzii* and *E. coli* from all the biopsies pooled together (Table 3) and by sample location (Table 4) was compared amongst patients with different intestinal disorders and healthy controls in order to determine whether or not their relative abundance could be employed as a useful indicator to distinguish between IBS and IBD patients, and within IBD phenotypes.

F. prausnitzii abundance

F. prausnitzii abundance decreased in IBD patients, especially CD patients ($P < 0.001$), whereas IBS patients more closely resembled the H group (Table 3). Within UC patients, those with proctitis and extensive UC presented intermediate *F. prausnitzii* levels between CD patients and H subjects. In CD patients, those with the lowest levels of this bacterium were CD patients with ileal involvement (either I-CD or IC-CD), and CD patients with stricturing disease behavior, whereas C-CD patients resembled UC. ROC curve analysis, applied to test the accuracy of the indicators to differentiate between two groups of patients, confirmed that the reduction of *F. prausnitzii* abundance is a good discriminator for IBD patients, when compared to the H subjects and, more interestingly, with IBS patients (Table 5). The specificity was also improved when proctitis patients were removed from the analysis. Moreover, this indicator accurately distinguished I-CD patients from UC patients and, more interestingly also from C-CD patients. Precisely, when comparing I-CD patients with C-CD, the AUC values were greater than 0.772, corresponding to 82.50% sensitivity and above 57.14% specificity at a set threshold (Table 5).

When analyzing data by sample location, the trend to distinguish these disease phenotypes was observed at rectum and colon level, although only statistical significance was reached for the latter (Table 4). In contrast, *F. prausnitzii* abundance in ileal samples was not a suitable indicator to distinguish between IBD phenotypes.

E. coli abundance

E. coli abundance varied differently in IBD subjects (Table 3). UC patients presented a reduced abundance of this species ($P = 0.002$), with the exception of those with extensive UC, which harbored similar abundances to the CD group. By comparison, CD patients showed increased levels of *E. coli* when compared to H and IBS patients. Within CD phenotypes, all reached statistical significance except for the IC-CD group, probably due to the high variability of the data. As regards IBS patients, this parameter only allowed their discrimination from CD patients as confirmed by ROC curve analysis (Table 5). Interestingly, ROC curve analysis also showed that *E. coli* might differentiate extensive UC and C-CD (Table 5), which are two different pathological entities that feature overlapping clinical manifestations and are therefore difficult to diagnose.

Analysis by location of the sample indicated that colonic *E. coli* quantification can be a good marker to differentiate these two IBD phenotypes (Table 4). It is of note that *E. coli* was approximately ten times more abundant in ulcerated biopsies of CD patients than in those taken from non-ulcerated zones [median values of $\log_{10}$ (16S rRNA gene copies/million bacterial 16S rRNA gene copies) from ulcerated ($N = 17$) 5.02 ± 0.88 and non-ulcerated zones ($N = 71$) 4.13 ± 1.07; $P = 0.009$].

Table 5
Area under the curve (AUC) obtained by receiver operating characteristic analysis (ROC curve) to establish the usefulness of *F. prausnitzii*, *E. coli* and the F-E index to distinguish amongst different intestinal disorders (H, controls; IBD, inflammatory Bowel Disease; IBS, Irritable Bowel Syndrome; UC, ulcerative colitis; CD, Crohn's disease; I-CD, ileal CD; IC-CD, ileocolonic CD, C-CD, colonic CD). Sensitivity and specificity values at a set threshold have been included for comparative purposes. Only analysis with AUC values above 0.6 are shown as a test is considered to be suitable if the AUC range from 0.6 to 0.75, and to have good sensitivity and specificity if the AUC range from 0.75 to 0.9.

	AUC	Sensitivity (%)	Specificity (%)
F. prausnitzii			
H vs IBD	0.765	81.35	55.17
H vs IBD (without proctitis patients)	0.778	81.35	61.44
IBS vs IBD	0.696	80.77	54.60
IBS vs IBD (without proctitis patients)	0.710	80.76	61.44
I-CD vs C-CD	0.772	82.50	57.14
I-CD vs UC	0.793	82.50	53.84
E. coli			
IBS vs CD	0.693	82.29	57.69
C-CD vs extensive UC	0.636	86.67	35.71
F-E index			
IBS vs CD	0.797	80.21	61.54
IBS vs I-CD	0.868	80.77	72.50
IBS vs IC-CD	0.746	80.76	52.00
IBS vs C-CD	0.784	80.76	57.14
C-CD vs extensive UC	0.767	80.00	60.71

F. prausnitzii-E. coli index

Although both bacterial species were confirmed to be good indicators of IBD dysbiosis, we further investigated if the discriminatory power was enhanced when analyzing both species together. Thus, an index was calculated subtracting *E. coli* numbers from *F. prausnitzii* abundance and this data was further normalized to total bacterial 16S rRNA gene copies and to human cell numbers to correct for variations due to sample size (as detailed in Materials and Methods section).

When all the biopsies from different locations were pooled for analysis, a positive F-E index, indicating a predominance of *F. prausnitzii* over *E. coli*, was observed in H, IBS and UC patients, suggesting that these three groups of patients were undistinguishable from each other (Table 3). Nevertheless, the differentiation of IBS patients from all the CD subjects, irrespective of their disease location and behavior, improved when the F-E index was used in spite of the bacterial indicators alone. Noteworthy, using the F–E index we gained sensitivity (80.00%) and specificity (60.71%) to differentiate extensive UC from C-CD patients, which was not possible

when considering *F. prausnitzii* alone and was achieved with low specificity (35.71%) when only taking *E. coli* into account (Table 5). Interestingly, negative values of the index were mainly reached for those CD patients with ileal involvement, indicating that in this subgroup of patients, *E. coli* populations numerically dominate that of *F. prausnitzii*.

When data was analyzed by sample location, all these features were observed in both rectum and colon samples, however the latter was shown to be the most discriminatory sample (Table 4). Conversely, from the ileum samples, the F-E index of IBS patients also reached negative values that hampered the differentiation with CD patients. Moreover, C-CD patients showed higher values of the F-E index, resembling UC patients. Thus, our data suggests that ileal samples alone are not suitable for a correct diagnosis. However, the higher F-E index for C-CD patients for ileal samples in comparison to that in IC-CD patients provides an additional discrimination point as the two disease phenotypes had similar F-E index from colon samples.

The usefulness of a ratio F/E (16S rRNA *F. prausnitzii* genes/16S rRNA *E. coli* genes) and $\log_{10}$ ratio F/E was also evaluated. All the indexes achieved similar scores concerning discrimination between disorders and disease phenotypes, but the accuracy of discrimination measured by the area under the ROC curve was better with the F-E index (data not shown).

Correlation between F. prausnitzii and E. coli abundances in healthy subjects, IBS and IBD patients by disease phenotype

F. prausnitzii and *E. coli* numbers were analyzed in order to determine if they were positively or negatively correlated, and whether this could provide supporting evidence about a putative common factor affecting negatively/positively both bacterial populations in a given patient or about a direct/indirect effect of one population over the other (Fig. 1).

In H subjects, *E. coli* abundance fluctuated over a 5-$\log_{10}$ span irrespective of *F. prausnitzii* quantity, which in turn was reasonably stable (2-$\log_{10}$ span) within this group of subjects. No correlation between these two species was found. Similar results were observed for IBS patients irrespective of disease phenotype. Interestingly, in UC patients there was a positive correlation between *F. prausnitzii* and *E. coli*. This could not be associated with an increase in total bacteria, as the abundance of both species was normalized to total bacterial 16S rRNA gene copies. Although this trend was observed for all UC phenotypes (data not shown), statistical significance was achieved only when patients with extensive UC were considered (Fig. 1). In CD patients *F. prausnitzii* quantity was extremely variable and was spread over a 6-$\log_{10}$ span, whereas *E. coli* abundances were as dispersed as in H subjects, but reaching higher values. Whereas no correlation was found when all the CD phenotypes were grouped, C-CD patients considered alone showed positive correlation resembling that observed in UC patients, although these did not reach statistical significance (Fig. 1). Moreover, a tendency that suggests a possible negative correlation between *F. prausnitzii* and *E. coli* was observed when analyzing those patients with ileal involvement (Fig. 1) with 21% of the patients with I-CD and 15% of those with IC-CD exhibiting an increase in *E. coli* abundance with a concomitant decrease in *F. prausnitzii* numbers in comparison to H subjects. This suggests that the microbial imbalance is not homogeneously distributed among all the patients and that some feature a more severe dysbiosis. It is of note that the trend of a negative correlation between *F. prausnitzii* and *E. coli* was stronger when patients under anti-tumor necrosis factor (TNF) α therapy were removed from the analysis ($\rho = -0.237$, $P = 0.105$).

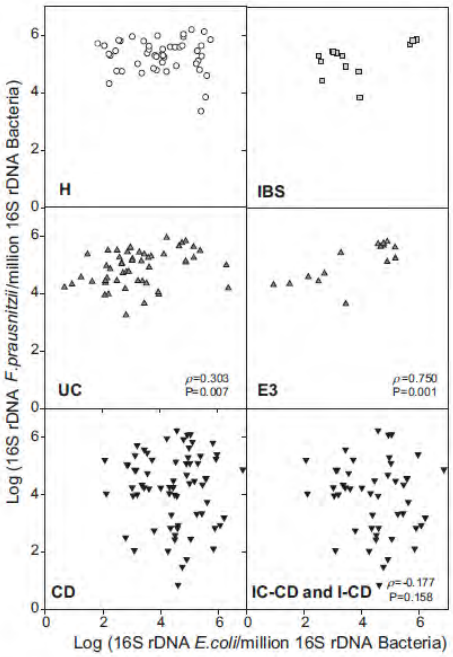


Fig. 1. Spearman correlation between mucosa-associated *F. prausnitzii* and *E. coli* in control (H), Irritable Bowel Syndrome (IBS), Ulcerative Colitis (UC), and Crohn's disease (CD) patients (16S rRNA gene copies/million bacterial 16 S rRNA gene copies). Correlations in extensive ulcerative colitis (E3) and CD patients with ileal involvement (I-CD and IC-CD) are specified.

F. prausnitzii and E. coli abundance in relation to patients clinical data

F. prausnitzii and *E. coli* abundances were compared between active and inactive patients (active CD and UC were defined by a CDAI of >150 (Best et al., 1976) and a Mayo score >3, respectively) in order to determine if these indicators vary according to the activity status of the patient. *F. prausnitzii* abundance did not differ between active and inactive IBD patients (Fig. 2A). Conversely, *E. coli* load was increased in active IBD patients, although only statistically significant differences were found for CD patients (Fig. 2B), and particularly in those with ileal involvement (Table S3). However, no correlation was found between *E. coli* abundance with the SES-CD, nor with the levels of blood C-reactive protein (data not shown).

We also investigated whether or not the abundance of *F. prausnitzii* and *E. coli* at the time of sampling could be correlated with time to recurrence of disease in five inactive CD patients (three with I-CD, one with IC-CD and one with C-CD) of whom we had available information on disease relapse (Fig. 3). Interestingly, *F. prausnitzii* abundance correlated positively with months to next flare-up ($\rho = 0.660$, $p < 0.001$) indicating that the higher the *F. prausnitzii* abundance, the longer the remission. In contrast, *E. coli* was negatively correlated with months to next flare-up ($\rho = -0.129$, $P = 0.030$), suggesting that when *E. coli* numbers are higher, the period of remission is shortened. These results suggest that the abundance of both species might be applicable as predictors of disease recurrence.

F. prausnitzii and *E. coli* quantities were also analyzed taking into account whether or not the patients required intestinal resection during the course of the disease. *F. prausnitzii* abundance was reduced in those CD patients that underwent intestinal resection [median values of $\log_{10}$ (16S rRNA gene copies/million bacterial 16S rRNA gene copies) from non-resected ($N = 30$) 4.57 ± 1.40, and

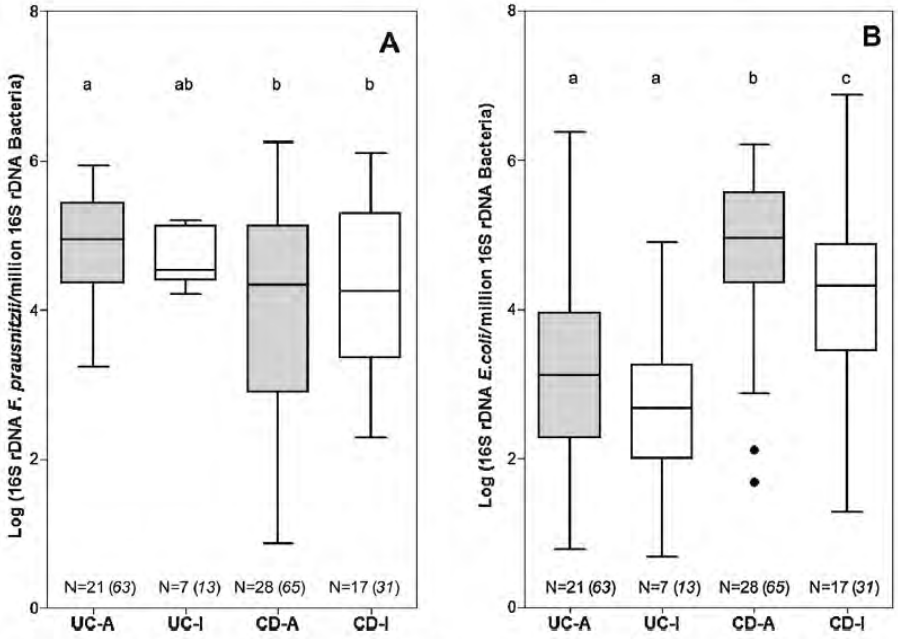


Fig. 2. *F. prausnitzii* (A) and *E. coli* (B) abundances categorized by activity status of Ulcerative Colitis (UC) and Crohn's disease (CD) patients (grey, active; white, inactive). The number of patients and biopsies (*values in italics*) in each group has been indicated. Homogeneous subgroups ($P > 0.05$) within each panel are indicated with the same superscript.

resected (N = 11) 3.95 ± 0.78; P = 0.009], whereas *E. coli* numbers were similar between resected and non-resected patients.

F. prausnitzii and E. coli abundances by treatment

In order to establish which therapy might have an effect in correcting dysbiosis, the abundance of both species was analyzed by current medication of the patients at the time of sampling.

All IBD patients regardless of their medication showed decreased *F. prausnitzii* loads when compared with the H group, indicating that this species abundance was not restored by any of the therapies considered in this study (Table 6). No differences in *F. prausnitzii* abundance were observed between medications within any disease phenotype. Conversely, *E. coli* numbers were lower in I-CD patients under anti-TNFα treatment, suggesting that this treatment has a direct effect on modulating the abundance of this pro-inflammatory bacterium in the gut of patients with this disease phenotype (Table 6).

As we observed that *E. coli* numbers were lower in CD patients under anti-TNFα treatment, we enrolled a subgroup of 10 CD patients (4 C-CD, 2 IC-CD and 4 I-CD) who started TNFα inhibition therapy with adalimumab in a follow-up study, who were

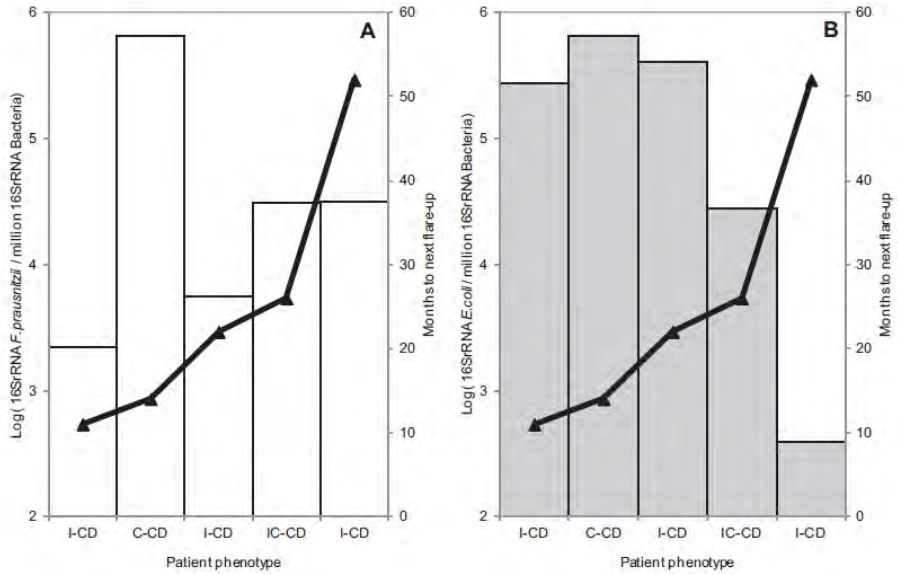


Fig. 3. Retrospective study to determine the usefulness of *F. prausnitzii* (A) and *E. coli* (B) abundances to predict time to flare-ups (black triangles) in CD patients. Disease phenotypes of the patients has been indicated (I-CD, ileal CD; IC-CD, ileocolonic CD; C-CD, colonic CD).

Table 6
F. prausnitzii and *E. coli* abundances in different Inflammatory Bowel Disease phenotypes by medication at sampling.

	*F. prausnitzii**§						*E. coli**					
	N patients (*n biopsies*)	No treatment or mesalazine	N patients (*n biopsies*)	Moderate immunosup-presants	N patients (*n biopsies*)	Anti-TNF	P value	N patients (*n biopsies*)	No treatment or mesalazine	N patients (*n biopsies*)	Moderate immunosup-presants	N patients (*n biopsies*)	Anti-TNF	P value
UC	**17 (38)**	**4.98 ± 0.59**	**4 (14)**	**4.42 ± 0.56**	**5 (14)**	**4.94 ± 0.80**	**ns**	**17 (47)**	**3.20 ± 1.28**	**4 (14)**	**2.43 ± 0.92**	**5 (14)**	**3.21 ± 0.91**	**ns**
E1	6 (18)	5.12 ± 0.31						6 (18)	3.04 ± 0.75					
E2	5 (9)	4.42 ± 0.41	4 (14)	4.42 ± 0.56	3 (12)	5.44 ± 0.77	ns	5 (9)	3.30 ± 2.06	4 (14)	2.43 ± 0.92	3 (12)	3.39 ± 0.71	ns
E3	4 (11)	5.62 ± 0.59			2 (2)	3.99 ± 0.47	ns	4 (11)	4.57 ± 1.31			2 (2)	2.20 ± 1.77	ns
CD	**17 (40)**	**4.57 ± 1.45**	**17 (30)**	**4.07 ± 0.96**	**11 (21)**	**4.04 ± 1.33**	**ns**	**17 (40)**	**4.32 ± 0.88**	**17 (30)**	**4.96 ± 1.18**	**11 (21)**	**4.50 ± 1.03**	**0.021**
C-CD	6 (13)	5.15 ± 0.67	4 (10)	4.27 ± 0.89	3 (5)	5.72 ± 1.46	ns	6 (13)	4.71 ± 0.71	4 (10)	4.73 ± 0.93	3 (5)	3.19 ± 0.90	ns
IC-CD	4 (10)	4.53 ± 0.96	6 (8)	4.30 ± 0.73	3 (6)	3.78 ± 1.57	ns	4 (10)	3.75 ± 0.69	6 (8)	4.14 ± 1.55	3 (6)	4.99 ± 0.98	ns
I-CD	6 (17)	4.21 ± 1.76	7 (12)	3.44 ± 1.02	5 (10)	3.97 ± 1.06	ns	6 (17)	4.26 ± 1.04	7 (12)	5.45 ± 0.75	5 (10)	4.51 ± 0.96	0.002

* Homogeneous subgroups ($P > 0.05$) within each variable (column) are indicated with the same superscript; TNF, tumour necrosis factor; ns, not significant.
Median $\log_{10}$ 16S rRNA gene copies/million bacterial 16S rRNA gene copies ± standard deviations

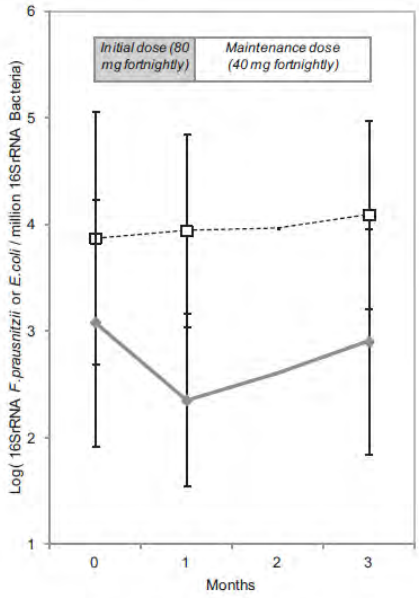


Fig. 4. *F. prausnitzii* (white squares) and *E. coli* (grey diamonds) abundances over a three months period in rectal biopsies of a group of patients who started adalimumab therapy.

monitored before starting the treatment and at months one and three after initiation. Although *F. prausnitzii* abundance did not increase substantially after adalimumab treatment, in agreement with the previous results *E. coli* numbers markedly decreased when adalimumab was given at a dose of 80 mg every two weeks during the first month (induction dose) and were maintained slightly lower than before treatment when the dose was decreased to 40 mg fortnightly (maintenance dose) (Fig. 4). This result was not statistically supported probably due to the low number of patients enrolled in this trial and the high variability between subjects.

Discussion

In the present study we have analyzed the abundance of mucosa associated *F. prausnitzii* and *E. coli* in H, IBS and IBD subjects, paying careful attention to the diversity of disease phenotypes and clinical features of the patients. We show that these two bacterial species can be good indicators to assist in IBD diagnostics and, for some disease phenotypes, in disease prognosis. Moreover, new information about which current therapies in IBD might correct dysbiosis towards "normobiosis" (Roberfroid et al., 2010) is also revealed.

Our data showed that *F. prausnitzii* and *E. coli* abundances behave differently among intestinal disorders and IBD phenotypes and confirmed quantitatively that *F. prausnitzii* is a specific IBD dysbiosis indicator that has allowed us to distinguish UC and CD patients from those with IBS. This is in agreement with previous work based on fecal samples (Swidsinski et al., 2008) although this study did not determine to what extent the bacterial imbalance found was a common feature of all the disease phenotypes. Our study confirmed that the depletion in *F. prausnitzii* abundance is a feature of all the IBD patients with the exception of those UC patients with proctitis and extensive UC, and therefore additional information is required in order to distinguish these disease phenotypes. Using *E. coli* as a second indicator in combination with *F. prausnitzii* we gained discrimination power and UC proctitis patients were distinguishable from H subjects.

When using *F. prausnitzii* or *E. coli* as single indicators it was not possible to fully distinguish within all the UC and CD phenotypes.

In contrast, the F-E index allowed a neat differentiation of I-CD patients with respect to other IBD subgroups, that could be useful to assist differential diagnosis between I-CD and IC-CD. More interestingly, the F–E index allowed for good differentiation of C-CD patients from those patients with extensive UC, as well as for the other UC disease phenotypes, which is of relevance for diagnostic purposes as these two disorders may present similar clinical manifestations (Bernstein et al., 2010; Nikolaus and Schreiber, 2007). The quantification of these two species could therefore be implemented as a reliable marker to aid diagnosis of these intestinal disorders. Unfortunately, it was not possible to distinguish distal UC from extensive UC, or C-CD patients from those with IC-CD by using these indicators or the F-E index, therefore, additional bacterial indicators are needed to properly distinguish all the IBD phenotypes. Further investigations to test the usefulness of the indicators to assign disease phenotype at early disease stages would be also of interest.

Our results suggest that the colon is the location that allows us to distinguish most of the phenotypes and therefore should be the location of choice to sample. Nevertheless, ileal samples could provide an additional discrimination point to support differentiation between certain disease phenotypes such as IC-CD and C-CD. Although dysbiosis observed in the rectum resembled that found in the colon, additional studies with larger number of rectal samples should be performed in order to corroborate this observation. Besides, testing the usefulness of the microbiological biomarkers presented here in non-invasive fecal samples would be of interest in order to assist in early diagnosis.

F. prausnitzii abundance was similar between active and inactive patients with the same IBD phenotype, which indicates that this species can be a reliable marker to screen IBD patients even in remission. Although our results do not concur with previous studies based on fecal samples (Duboc et al., 2013; Sokol et al., 2009), a reduction in *F. prausnitzii* numbers in CD patients in remission has already been reported in studies based on biopsies (Willing et al., 2009). We hypothesize that the depletion in *F. prausnitzii* at the mucosal level (which is the site of microbial recognition by the host and where the inflammatory process is developing) may be more evident than in feces of patients in remission.

In contrast, *E. coli* abundance was higher in active CD patients by comparison with those in remission at sampling, which supports the hypothesis that *E. coli* is involved in CD pathogenesis (Darfeuille-Michaud et al., 2004; Martin et al., 2004; Martinez-Medina et al., 2009; Sasaki et al., 2007). It is of note that indices of endoscopic activity (SES-CD) and general inflammation (C-reactive protein) did not correlate with imbalances in these indicators and, reinforces the necessity of using several parameters to define a real deep remission. It may therefore be worth considering to assess "microbiological remission" as a new parameter in the future.

In agreement with a previous study (Sokol et al., 2008) lower numbers of *F. prausnitzii* were observed in resected CD patients although our study did not allow us to decipher whether this depletion could be associated with the need for surgical intervention. Thus there should be further investigation to assess the usefulness of this biomarker to precisely predict when such intervention might be needed.

Concerning the applicability of these two indicators for prognostic purposes, we observed that increased levels of *E. coli* were associated with a relapse in a short period of time in CD patients, whereas high levels of *F. prausnitzii* and low levels of *E. coli* were associated with longer remission periods. Our data is in agreement with the previous work of Sokol et al. (2008) reporting that a reduction in *F. prausnitzii* abundance was associated with endoscopic recurrence of the disease (Sokol et al., 2008, 2009). Nevertheless, we observed that high *F. prausnitzii* abundance without a decrease in *E. coli* numbers did not ensure a long remission period, therefore the subgroup of patients analyzed, predominated by I-CD patients, showed that an imbalance in *E. coli* abundance plays a greater role in inducing inflammation than the depletion of the *F. prausnitzii* load. This suggests that *F. prausnitzii* and *E. coli* are potentially useful for prognostics in I-CD. However further prospective studies in a larger cohort of patients are needed to confirm this hypothesis.

Interestingly, in this study we observed that a correlation exists between the abundance of these two species in IBD patients, a feature that, to our knowledge, has not been described to date. In UC patients, the relative abundance of *F. prausnitzii* and *E. coli* were positively correlated, suggesting that under this intestinal disorder populations of both species might be affected similarly by gut environment or host factors. Conversely, a negative correlation trend was observed in CD patients with ileal disease location, with *E. coli* being more abundant than *F. prausnitzii*. This negative correlation between species specially associated to I-CD patients leads us to hypothesize that both species are directly linked to the disease pathogenesis by playing different roles. This hypothesis is sustained by several reports that implicate the adherent-invasive *E. coli* (AIEC) pathovar in CD pathogenesis (Darfeuille-Michaud et al., 2004; Martin et al., 2004; Martinez-Medina et al., 2009; Sasaki et al., 2007) and those that postulate that a reduction of *F. prausnitzii* might be a crucial factor to enhance disease recurrence (Sokol et al., 2008, 2009). However, we could not confirm whether or not the observed increase in *E. coli* was due to the AIEC pathovar since to date, no molecular tool for its specific quantification is available. Another possibility to explain the negative correlation between the two species is that changes in gut or host environmental factors may be implicated. For instance, bile salts, whose composition has been recently demonstrated to be altered in IBD patients (Duboc et al., 2013), can negatively affect *F. prausnitzii* growth (Lopez-Siles et al., 2012) and also induce the expression of virulence factors in *E. coli* (Chassaing et al., 2013). Moreover, a direct or indirect effect of one population on the other also cannot be ruled out, and further co-culture experiments would be helpful to fully elucidate the interactions between these two species.

Our results give valuable insight as to how current therapies applied in IBD treatment might be leading to a correction of dysbiosis by modulating the populations of these two species. We observed that *E. coli* numbers were lower in I-CD patients under anti-TNFα treatment when compared with other therapies, and it was further corroborated in a prospective study in which CD patients were treated with adalimumab. It has been previously reported that TNFα promotes the expression of carcinoembryonic antigen-related cell adhesion molecule 6, which is a molecule used by *E. coli* to adhere to enterocytes via the interaction with type 1 pili (Barnich et al., 2007). Besides, AIEC strains have been reported to be more efficient than non-AIEC strains isolated from the intestinal mucosa of IBD patients and controls, at colonizing the gut due to special mutations in the FimH adhesion of type 1 pili (Dreux et al., 2013). We hypothesize that the blockage of TNFα can lead to lower expression of carcinoembryonic antigen-related cell adhesion molecule 6 which in turn might imply lower AIEC colonization. However, to prove this hypothesis specific quantification of this pathovar is needed. In contrast, none of the current medication regimes analyzed in the present study was shown to be effective in restoring the *F. prausnitzii* populations. Therefore, it is probable that to restore this species it might be necessary to re-establish the overall ecological conditions in the gut environment.

Conclusions

Our study confirms that *F. prausnitzii* and *E. coli* are good indicators of IBD dysbiosis and provides evidence for the applicability for disease diagnostics allowing the differentiation of IBD from IBS

and also between some IBD subtypes as C-CD from extensive UC. We further investigated the potential applicability for prognostics and, our data, although preliminary, allows us to conclude that this tool could be used as a supporting prognostic tool in CD patients since the remission in I-CD patients was associated with the abundance of these two species. The present study shows that current therapies are not sufficient to counterbalance dysbiosis and further investigations are required to show which other factors, other than medication, might help to revert bacterial populations back to a typical structure.

Acknowledgements

This work was partially funded by the Spanish Ministry of Education and Science through project SAF2010-15896. Mireia Lopez-Siles was recipient of an FI grant from the Generalitat de Catalunya (2010FI_B2 00135), which receives support from the European Union Commissionate. Prof. Harry J. Flint and Dr. Sylvia H. Duncan acknowledge support from the Scottish Government Food, Land and People programme. We thank Ms. Natàlia Adell from the Serveis Tècnics de Recerca for statistical assistance. We are grateful to Dr. Laia Calvó (Research Unit, Institut d'Assistència Sanitària, Salt, Spain) for her assistance in qPCR design and to Dr. Rosalia Trias (Universitat de Girona, Spain), who critically revised the manuscript. We appreciate the generosity of the patients who freely gave their time and samples to make this study possible, and the theatre staff of all centers for their dedication and careful sample collection.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ijmm.2014.02.009.

References

Ahmed, S., Macfarlane, G.T., Fite, A., McBain, A.J., Gilbert, P., Macfarlane, S., 2007. Mucosa-associated bacterial diversity in relation to human terminal ileum and colonic biopsy samples. Appl. Environ. Microbiol. 73, 7435–7442.

Barcenilla, A., Pryde, S.E., Martin, J.C., Duncan, S.H., Stewart, C.S., Henderson, C., Flint, H.J., 2000. Phylogenetic relationships of butyrate-producing bacteria from the human gut. Appl. Environ. Microbiol. 66, 1654–1661.

Barnich, N., Carvalho, F., Glasser, A.L., Darcha, C., Jantscheff, P., Allez, M., Peeters, H., Bommelaer, G., Desreumaux, P., Colombel, J.F., Darfeuille-Michaud, A., 2007. CEACAM6 acts as a receptor for adherent-invasive *E. coli*, supporting ileal mucosa colonization in Crohn disease. J. Clin. Invest. 117, 1566–1574.

Baumgart, D.C., Carding, S.R., 2007. Inflammatory bowel disease: cause and immunobiology. Lancet 369, 1627–1640.

Baumgart, D.C., Sandborn, W.J., 2007. Inflammatory bowel disease: clinical aspects and established and evolving therapies. Lancet 369, 1641–1657.

Bernstein, C.N., Shanahan, F., 2008. Disorders of a modern lifestyle: reconciling the epidemiology of inflammatory bowel diseases. Gut 57, 1185–1191.

Bernstein, C.N., Fried, M., Krabshuis, J.H., Cohen, H., Eliakim, R., Fedail, S., Gearry, R., Goh, K.L., Hamid, S., Khan, A.G., LeMair, A.W., Malfertheiner Ouyang, Q., Rey, J.F., Sood, A., Steinwurz, F., Thomsen, O.O., Thomson, A., Watermeyer, G., 2010. World Gastroenterology Organization Practice Guidelines for the diagnosis and management of IBD in 2010. Inflamm. Bowel Dis. 16, 112–124.

Best, W.R., Becktel, J.M., Singleton, J.W., Kern Jr., F., 1976. Development of a Crohn's disease activity index. National Cooperative Crohn's Disease Study. Gastroenterology 70, 439–444.

Bustin, S.A., Benes, V., Garson, J.A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M.W., Shipley, G.L., Vandesompele, J., Wittwer, C.T., 2009. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. Clin. Chem. 55, 611–622.

Cato, E.P., Salmon, C.W., Moore, W.E.C., 1974. *Fusobacterium prausnitzii* (Hauduroy et al.) Moore and Holdeman: emended description and designation of neotype strain. Int. J. Syst. Bacteriol. 24, 225–229.

Chassaing, B., Etienne-Mesmin, L., Bonnet, R., Darfeuille-Michaud, A., 2013. Bile salts induce long polar fimbriae expression favouring Crohn's disease-associated adherent-invasive *Escherichia coli* interaction with Peyer's patches. Environ. Microbiol. 15, 355–371.

Darfeuille-Michaud, A., Boudeau, J., Bulois, P., Neut, C., Glasser, A.L., Barnich, N., Bringer, M.A., Swidsinski, A., Beaugerie, L., Colombel, J.F., 2004. High prevalence of adherent-invasive *Escherichia coli* associated with ileal mucosa in Crohn's disease. Gastroenterology 127, 412–421.

De Cruz, P., Prideaux, L., Wagner, J., Ng, S.C., McSweeney, C., Kirkwood, C., Morrison, M., Kamm, M.A., 2012. Characterization of the gastrointestinal microbiota in health and inflammatory bowel disease. Inflamm. Bowel Dis. 18, 372–390.

Dreux, N., Denizot, J., Martinez-Medina, M., Mellmann, A., Billig, M., Kisiela, D., Chattopadhyay, S., Sokurenko, E., Neut, C., Gower-Rousseau, C., Colombel, J.F., Bonnet, R., Darfeuille-Michaud, A., Barnich, N., 2013. Point mutations in FimH adhesin of Crohn's disease-associated adherent-invasive *Escherichia coli* enhance intestinal inflammatory response. PLoS Pathog 9, e1003141.

Duboc, H., Rajca, S., Rainteau, D., Benarous, D., Maubert, M.A., Quervain, E., Thomas, G., Barbu, V., Humbert, L., Despras, G., Bridonneau, C., Dumetz, F., Grill, J.P., Masliah, J., Beaugerie, L., Cosnes, J., Chazouilleres, O., Poupon, R., Wolf, C., Mallet, J.M., Langella, P., Trugnan, G., Sokol, H., Seksik, P., 2013. Connecting dysbiosis, bile-acid dysmetabolism and gut inflammation in inflammatory bowel diseases. Gut 62, 531–539.

Duncan, S.H., Hold, G.L., Harmsen, H.J., Stewart, C.S., Flint, H.J., 2002. Growth requirements and fermentation products of *Fusobacterium prausnitzii*, and a proposal to reclassify it as *Faecalibacterium prausnitzii* gen. nov., comb. nov. Int. J. Syst. Evol. Microbiol. 52, 2141–2146.

Elson, C.O., Cong, Y., 2012. Host-microbiota interactions in inflammatory bowel disease. Gut Microbes 3, 332–344.

Frank, D.N., St Amand, A.L., Feldman, R.A., Boedeker, E.C., Harpaz, N., Pace, N.R., 2007. Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases. Proc. Natl. Acad. Sci. U.S.A. 104, 13780–13785.

Furet, J.P., Firmesse, O., Gourmelon, M., Bridonneau, C., Tap, J., Mondot, S., Doré, J., Corthier, G., 2009. Comparative assessment of human and farm animal faecal microbiota using real-time quantitative PCR. FEMS Microbiol. Ecol. 68, 351–362.

Huijsdens, X.W., Linskens, R.K., Mak, M., Meuwissen, S.G.M., Vandenbroucke-Grauls, C.M.J.E., Savelkoul, P.H.M., 2002. Quantification of bacteria adherent to gastrointestinal mucosa by real-time PCR. J. Clin. Microbiol. 40, 4423–4427.

Joossens, M., Huys, G., Cnockaert, M., De Preter, V., Verbeke, K., Rutgeerts, P., Vandamme, P., Vermeire, S., 2011. Dysbiosis of the faecal microbiota in patients with Crohn's disease and their unaffected relatives. Gut 60, 631–637.

Kaser, A., Zeissig, S., Blumberg, R.S., 2010. Inflammatory bowel disease. Annu. Rev. Immunol. 28, 573–621.

Kleessen, B., Kroesen, A.J., Buhr, H.J., Blaut, M., 2002. Mucosal and invading bacteria in patients with inflammatory bowel disease compared with controls. Scand. J. Gastroenterol. 37, 1034–1041.

Lane, D.J., 1991. In: Stackebrandt, E., Goodfellow, M. (Eds.), 16S/23S rRNA Sequencing. John Willy and Sons, New York.

Lopez-Siles, M., Khan, T.M., Duncan, S.H., Harmsen, H.J., Garcia-Gil, L.J., Flint, H.J., 2012. Cultured representatives of two major phylogroups of human colonic *Faecalibacterium prausnitzii* can utilize pectin, uronic acids, and host-derived substrates for growth. Appl. Environ. Microbiol. 78, 420–428.

Louis, E., Collard, A., Oger, A.F., Degroote, E., El Yafi, F.A.N., Belaiche, J., 2001. Behaviour of Crohn's disease according to the Vienna classification: changing pattern over the course of the disease. Gut 49, 777–782.

Louis, P., Duncan, S.H., McCrae, S.I., Millar, J., Jackson, M.S., Flint, H.J., 2004. Restricted distribution of the butyrate kinase pathway among butyrate-producing bacteria from the human colon. J. Bacteriol. 186, 2099–2106.

Manichanh, C., Borruel, N., Casellas, F., Guarner, F., 2012. The gut microbiota in IBD. Nat Rev Gastroenterol. Hepatol. 9, 599–608.

Martin, H.M., Campbell, B.J., Hart, C.A., Mpofu, C., Nayar, M., Singh, R., Englyst, H., Williams, H.F., Rhodes, J.M., 2004. Enhanced *Escherichia coli* adherence and invasion in Crohn's disease and colon cancer. Gastroenterology 127, 80–93.

Martinez-Medina, M., Aldeguer, X., Gonzalez-Huix, F., Acero, D., Garcia-Gil, L.J., 2006. Abnormal microbiota composition in the ileocolonic mucosa of Crohn's disease patients as revealed by polymerase chain reaction-denaturing gradient gel electrophoresis. Inflamm. Bowel Dis. 12, 1136–1145.

Martinez-Medina, M., Aldeguer, X., Lopez-Siles, M., Gonzalez-Huix, F., Lopez-Oliu, C., Dahbi, G., Blanco, J.E., Blanco, J., Garcia-Gil, L.J., Darfeuille-Michaud, A., 2009. Molecular diversity of *Escherichia coli* in the human gut: new ecological evidence supporting the role of adherent-invasive *E. coli* (AIEC) in Crohn's disease. Inflamm. Bowel Dis. 15, 872–882.

Miquel, S., Martin, R., Rossi, O., Bermudez-Humaran, L., Chatel, J., Sokol, H., Thomas, M., Wells, J., Langella, P., 2013. *Faecalibacterium prausnitzii* and human intestinal health. Curr. Opin. Microbiol. 16, 255–261.

Mondot, S., Kang, S., Furet, J.P., Aguirre de Carcer, D., McSweeney, C., Morrison, M., Marteau, P., Dore, J., Leclerc, M., 2011. Highlighting new phylogenetic specificities of Crohn's disease microbiota. Inflamm. Bowel Dis. 17, 185–192.

Nikolaus, S., Schreiber, S., 2007. Diagnostics of inflammatory bowel disease. Gastroenterology 133, 1670–1689.

Pineton de Chambrun, G., Peyrin-Biroulet, L., Lemann, M., Colombel, J.F., 2010. Clinical implications of mucosal healing for the management of IBD. Nat. Rev. Gastroenterol. Hepatol. 7, 15–29.

Roberfroid, M., Gibson, G.R., Hoyles, L., McCartney, A.L., Rastall, R., Rowland, I., Wolvers, D., Watzl, B., Szajewska, H., Stahl, B., Guarner, F., Respondek, F., Whelan, K., Coxam, V., Davicco, M.J., Leotoing, L., Wittrant, Y., Delzenne, N.M., Cani, P.D., Neyrinck, A.M., Meheust, A., 2010. Prebiotic effects: metabolic and health benefits. Br. J. Nutr. 104 (Suppl 2), S1L 63.

Sartor, R.B., 2006. Mechanisms of disease: pathogenesis of Crohn's disease and ulcerative colitis. Nat. Clin. Pract. Gastroenterol. Hepatol. 3, 390–407.

Sasaki, M., Sitaraman, S.V., Babbin, B.A., Gerner-Smidt, P., Ribot, E.M., Garrett, N., Alpern, J.A., Akyildiz, A., Theiss, A.L., Nusrat, A., Klapproth, J.M., 2007. Invasive *Escherichia coli* are a feature of Crohn's disease. Lab. Invest. 87, 1042–1054.

Schultsz, C., Van Den Berg, F.M., Ten Kate, F.W., Tytgat, G.N., Dankert, J., 1999. The intestinal mucus layer from patients with inflammatory bowel disease harbors high numbers of bacteria compared with controls. Gastroenterology 117, 1089–1097.

Seksik, P., Rigottier-Gois, L., Gramet, G., Sutren, M., Pochart, P., Marteau, P., Jian, R., Doré, J., 2003. Alterations of the dominant faecal bacterial groups in patients with Crohn's disease of the colon. Gut 52, 237–242.

Silverberg, M.S., Satsangi, J., Ahmad, T., Arnott, I.D., Bernstein, C.N., Brant, S.R., Caprilli, R., Colombel, J.F., Gasche, C., Geboes, K., Jewell, D.P., Karban, A., Loftus Jr., E.V., Pena, A.S., Riddell, R.H., Sachar, D.B., Schreiber, S., Steinhart, A.H., Targan, S.R., Vermeire, S., Warren, B.F., 2005. Toward an integrated clinical, molecular and serological classification of inflammatory bowel disease: Report of a Working Party of the 2005 Montreal World Congress of Gastroenterology. Can. J. Gastroenterol. 19 (Suppl A), 5–36.

Sokol, H., Pigneur, B., Watterlot, L., Lakhdari, O., Bermudez-Humaran, L.G., Gratadoux, J.J., Blugeon, S., Bridonneau, C., Furet, J.P., Corthier, G., Grangette, C., Vasquez, N., Pochart, P., Trugnan, G., Thomas, G., Blottiere, H.M., Dore, J., Marteau, P., Seksik, P., Langella, P., 2008. *Faecalibacterium prausnitzii* is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. Proc. Natl. Acad. Sci. U.S.A. 105, 16731–16736.

Sokol, H., Seksik, P., Furet, J.P., Firmesse, O., Nion-Larmurier, I., Beaugerie, L., Cosnes, J., Corthier, G., Marteau, P., Dore, J., 2009. Low counts of *Faecalibacterium prausnitzii* in colitis microbiota. Inflamm. Bowel Dis. 15, 1183–1189.

Suzuki, M.T., Taylor, L.T., DeLong, E.F., 2000. Quantitative analysis of small-subunit rRNA genes in mixed microbial populations via 5′-nuclease assays. Appl. Environ. Microbiol. 66, 4605–4614.

Swidsinski, A., Ladhoff, A., Pernthaler, A., Swidsinski, S., Loening-Baucke, V., Ortner, M., Weber, J., Hoffmann, U., Schreiber, S., Dietel, M., Lochs, H., 2002. Mucosal flora in inflammatory bowel disease. Gastroenterology 122, 44–54.

Swidsinski, A., Loening-Baucke, V., Vaneechoutte, M., Doerffel, Y., 2008. Active Crohn's disease and ulcerative colitis can be specifically diagnosed and monitored based on the biostructure of the fecal flora. Inflamm. Bowel Dis. 14, 147–161.

Weisburg, W.G., Barns, S.M., Pelletier, D.A., Lane, D.J., 1991. 16S ribosomal DNA amplification for phylogenetic study. J. Bacteriol. 173, 697–703.

Willing, B., Halfvarson, J., Dicksved, J., Rosenquist, M., Jarnerot, G., Engstrand, L., Tysk, C., Jansson, J.K., 2009. Twin studies reveal specific imbalances in the mucosa-associated microbiota of patients with ileal Crohn's disease. Inflamm. Bowel Dis. 15, 653–660.

Xavier, R.J., Podolsky, D.K., 2007. Unravelling the pathogenesis of inflammatory bowel disease. Nature 448, 427–434.

Yantiss, R.K., Odze, R.D., 2006. Diagnostic difficulties in inflammatory bowel disease pathology. Histopathology 48, 116–132.

Article IV

Changes in the abundance of mucosa-associated *Faecalibacterium prausnitzii* phylogroups I and II in the intestinal mucosa of inflammatory bowel disease and patients with colorectal cancer

Mireia Lopez-Siles[1], Margarita Martinez-Medina[1], Romà Surís-Valls[1], Xavier Aldeguer[2], Miriam Sabat-Mir[3], Sylvia H. Duncan[4], Harry J. Flint[4], L. Jesús Garcia-Gil[1]✉.

[1] Laboratory of Molecular Microbiology, Biology Departament, Universitat de Girona, Girona, Spain. [2] Department of Gastroenterology, Hospital Dr. Josep Trueta, Girona, Spain. [3] Department of Gastroenterology, Hospital Santa Caterina, Salt, Girona, Spain. [4] Microbial Ecology Group, Rowett Institute of Nutrition and Health, University of Aberdeen, Aberdeen, United Kingdom. ✉ Corresponding author, e-mail: jesus.garcia@udg.edu

Accepted in Inflammatory Bowel Disease on 3rd August 2015, doi:10.1097/MIB.0000000000000590

Abstract

Background

Faecalibacterium prausnitzii comprises 2 phylogroups, whose abundance in healthy and diseased gut and in conjunction with *Escherichia coli* has not yet been studied. This work aims to determine the contribution of *F. prausnitzii* phylogroup I and II in intestinal disease and to assess their potential diagnostic usefulness as biomarkers for gut diseases.

Methods

Total *F. prausnitzii*, its phylogroups and *E. coli* loads were determined by quantitative polymerase chain reaction targeting the 16S rRNA gene on biopsies from 31 healthy controls (H), 45 patients with Crohn's disease (CD), 25 patients with ulcerative colitis (UC), 10 patients with irritable bowel syndrome (IBS), and 20 patients with colorectal cancer (CRC). Data were normalized to total bacterial counts, and analyzed according to patients' disease location and clinical characteristics.

Results

Lower levels of both total *F. prausnitzii* and phylogroup I were found in subjects with CD, UC and CRC ($P<0.001$) compared with H subjects. Phylogroup I load was a better biomarker than total *F. prausnitzii* to discriminate subjects with gut disorders from H. Phylogroup II depletion was observed only in patients with CD ($P<0.001$) and can be potentially applied to differentiate ulcerative pancolitis from colonic CD. No statistically significant correlation between *E. coli* and any of the 2 *F. prausnitzii* phylogroups was found

in any group of patients or by IBD location. Phylogroup I was lower in active patients with whereas those CD with intestinal resection showed a reduction in phylogroup II. Treatments with mesalazine and immunosupressants did not result in the recovery of *F. prausnitzii* phylogroups abundance.

Conclusion

F. prausnitzii phylogroup I was depleted in CD, UC and CRC, whereas phylogroup II was specifically reduced in CD. Quantification of *F. prausnitzii* phylogroups may help to identify gut disorders, and to classify inflammatory bowel disease location.

Keywords

Faecalibacterium prausnitzii phylogroups, Crohn's disease, ulcerative colitis, colorectal cancer, irritable bowel syndrome.

Introduction

Faecalibacterium prausnitzii (Ruminococcaceae) is one of the 3 most abundant species found in the gut, representing between 2% and 20% of the fecal microbiota in healthy individuals, according to diversity studies of the human gut microbiome based on 16S rRNA gene analysis (1-6). This species has been reported to represent 6% of bacteria in mucosa-associated microbial communities (7), although some studies have indicated that these values can increase to around 20% to 50% in some individuals (8, 9).

In recent years, there has been increasing interest in *F. prausnitzii* given its potentially important role in promoting gut health (10, 11) through the formation of anti-inflammatory compounds (10-14) and enhancement of intestinal barrier function (15, 16). Many studies have shown that *F. prausnitzii* prevalence and abundance are reduced in different intestinal disorders (for review see Ref. (17) and references therein), although the depletion in *F. prausnitzii* numbers has been most extensively reported in inflammatory bowel disease (IBD). Low counts of this species have been observed in both fecal and mucosa-associated communities of adult patients with Crohn's disease (CD) (11, 18-21). Variable populations have been reported in patients with ulcerative colitis (UC) (7, 18, 19, 22-27), although the reduction of Firmicutes has been repeatedly observed in this disorder (25, 28, 29). A recent study conducted on 127 UC subjects points out that a reduction in *F. prausnitzii* is also involved in UC dysbiosis (25). Interestingly, lower counts of *Faecalibacterium*-related bacteria have also been observed in functional gut disorders such as irritable bowel syndrome (IBS) of alternating type (30), that in turn shares some features with IBD patients (31, 32), and in more severe intestinal disorders as colorectal cancer (CRC)(33). Taken together these

findings suggest that shifts in *F. prausnitzii* numbers occur under several pathological disorders, but it still remains to be established if this reduction is equivalent among different conditions, as few studies have considered several gut pathologies simultaneously.

Furthermore, relatively few studies have paid attention to the diversity within the genus *Faecalibacterium*. Recent phylogenetic analysis showed that mainly 2 different *F. prausnitzii* phylogroups, which include the current cultured representatives, are found in fecal samples of healthy subjects (14), but no data about the abundance of these phylogroups in gut disorders have been reported to date.

Many studies have reported that in addition to *F. prausnitzii* depletion, CD dysbiosis is characterized by an increase in *Escherichia coli* abundance, predominantly in patients with CD and ileal involvement (21, 34-36). A possible negative correlation between *F. prausnitzii* and *E. coli* has been observed in ileal CD (I-CD) patients (18), suggesting a direct/indirect effect of one population over the other. However it remains to be established whether or not this affects both phylogroups of *F. prausnitzii.*

This work is aimed at determining the variation of mucosa-associated *F. prausnitzii* phylogroups between healthy subjects and patients having several intestinal disorders to establish whether the imbalance in *F. prausnitzii* includes the overall population or specifically affects a particular phylogroup. Besides correlation between *F. prausnitzii* phylogroups and *E. coli* loads has also been analyzed. The prevalence and abundance of mucosa-associated *F. prausnitzii* and both phylogroups were determined in samples of patients with CD, UC, IBS and CRC and in healthy controls (H) at different locations of the gut. To this end, a novel multiplex quantitative polymerase chain reaction (qPCR) assay was developed for the quantification of the 2 known phylogroups within this species. Data were analyzed taking into account patients' most relevant clinical characteristics, to determine its usefulness to differentially diagnose patients with IBD and monitor the evolution of the disease. Medication at sampling was also considered to determine whether any of the current therapies are effective in correcting this species imbalance.

Materials and Methods

Patients, clinical data and sampling

A Spanish cohort consisting of 70 patients with IBD (45 CD and 25 UC), 10 patients with IBS, 20 patients with CRC, and 31 H was enrolled (Table 1). Subjects were recruited by the Gastroenterology Services of the Hospital Universitari Dr. Josep Trueta (Girona, Spain) and the Hospital Santa Caterina (Salt, Spain). Subjects were gender matched for all the groups. Concerning age, patients with CD were younger than those in

Table 1. Sample size and clinical characteristics of subjects.

	Healthy*	Irritable bowel syndrome	IBD		Colorectal cancer	p value§
			Crohn's disease	Ulcerative colitis		
N (patients)	31	10	45	25	20	
Age (mean years ± SD)	48.1±16.3	42.4±11.4	33.5±11.1	40.1±15.8	58.6±7.52	<0.001‡
Male (N, %)	16 (51.6%)	2 (20.0%)	26 (57.7%)	16 (64.0%)	14 (70.0%)	0.605†
Active (N, %)	na	na	28 (62.2%)	20 (80.0%)	na	0.059†
Previous surgery (N, %)	0	nd	9 (20.0%)	1 (4.0%)	nd	0.049†
Smokers (N, %)	0	0	8 (17.8%)	2 (8.0%)	5 (25.0%)	0.327†
Treatment (N, %) **						0.087†
No treatment			12 (26.7%)	13 (52.0%)		
Mesalazine	na	na	3 (6.7%)	3 (12.0%)	na	
Moderate immunosuppressant	na	na	16 (35.5%)	3 (12.0%)	na	
Anti-TNFα (infliximab, adalimumab)	na	na	10 (22.2%)	4 (16.0%)	na	
CD Montreal classification						
Age of diagnosis (N, %) **						0.257‡
diag < 16y (A1)	na	na	5 (11.1%)	1 (4.0%)	nd	
diag 17-40y (A2)	na	na	33 (73.3%)	13 (52.0%)	nd	
diag >41y (A3)	na	na	5 (11.1%)	8 (32.0%)	nd	
Location (N, %)						na
Ileal-CD (L1)	na	na	19 (42.2%)	na	na	
Colonic-CD (L2)	na	na	11 (24.4%)	na	na	
Ileocolonic-CD (L3)	na	na	14 (31.1%)	na	na	
Behavior (N, %) **						na
Non-stricturing, non-penetrating (B1)	na	na	30 (66.7%)	na	na	
Stricturing (B2)	na	na	9 (20.0%)	na	na	
UC classification (N, %) **						na
Ulcerative proctitis (E1)	na	na	na	6 (24.0%)	na	
Distal UC (E2)	na	na	na	11 (44.0%)	na	
Extensive UC or ulcerative pancolitis (E3)	na	na	na	6 (24.0%)	na	
IBS subtype (N, %) **						na
Diarrhea predominant type	na	2 (20.0%)	na	na	na	
Constipation predominant type	na	2 (20.0%)	na	na	na	
CRC subtype (N, %) **						na
Sporadic	na	na	na	na	14 (70.0%)	
Hereditary***	na	na	na	na	3 (15.0%)	

IBD, Inflammatory bowel disease; IBS, Irritable bowel syndrome; CRC, colorectal cancer; TNF, tumor necrosis factor; nd, not determined; na, not applicable

*Controls consisted of subjects who underwent colonoscopy for different reasons: 9/31 rectorrhagia, 11/31 colorectal cancer familial history and 11/31 abdominal pain.

** Medical treatment at the time of sampling was available in 41/45 CD patients, and 23/25 UC patients; Age of disease onset was available for 43/45 CD patients, and 22/25 UC patients; Disease behavior at last follow-up before the time of sampling was available in 39/45 CD patients, and none had penetrating CD (B3); Maximal disease extent at the time of sampling was available in 23/25 UC patients; disease subtype was available in 4/10 Irritable bowel syndrome patients, and none had alternating predominant type; presence or absence of relatives with CRC could only be clearly tracked in 17/20 CRC patients.

***Patients were included within this category if a first grade relative has had also CRC.

§ Groups were compared by non-parametric statistical tests, and p value ≤0.05 was considered significant; † χ^2 test; ‡ Mann-Whitney U test

the H group (P<0.001), whereas patients with CRC were significantly older than all the other groups (P≤0.019). Patients with IBD were diagnosed according to standard clinical, pathological and endoscopic criteria and categorized according to the Montreal classification (37). Patients with IBS were diagnosed according to Rome III criteria (available at <http://www.romecriteria.org/criteria/>). The diagnosis of CRC was established by colonoscopy and biopsy, and data correlated with high risk of developing this disease were recorded. The control group consisted of subjects undergoing colonoscopy for different reasons as rectorrhagia (N=9), CRC familial history (N=11), and abdominal pain (N=11). Clinically relevant data of all the patients were collected. None of the subjects received antimicrobial treatment for at least 2 months before colonoscopy.

Before colonoscopy, patients were subjected to cleansing of the gastrointestinal tract using Casenglicol after manufacturer's guidelines. During routine endoscopy, up to 3 biopsy samples per patient were taken from different locations along the gut (distal ileum, colon, and rectum) after standard procedures. All biopsies were immediately placed in sterile tubes without any buffer and stored at -80 ºC after completion of the whole endoscopic procedure and on analysis

Sample treatment and DNA extraction

After DNA extraction, biopsies were subjected to 2 mild ultrasound wash cycles to discard transient and loosely attached bacteria as previously reported (34). DNA was extracted using the NucleoSpin Tissue Kit (Macherey-Nagel GmbH &Co., Duren, Germany). The support protocol for gram-positive bacteria and the RNAse treatment step were performed. Genomic DNA was eluted with 10mM Tris-HCl (pH 7.4) and stored at –80 ºC until use. DNA concentration and purity of the extracts were determined with a NanoDrop ND-100 spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA).

Primers and hydrolysis probes design, and set up of a qPCR assay for *F. prausnitzii* phylogroups

To simultaneously quantify both *F. prausnitzii* phylogroups, a qPCR assay consisting of a unique pair of species-specific primers for *F. prausnitzii* and 2 hydrolysis probes targeting each *F. prausnitzii* phylogroup was designed.

Sequences of the 16S rRNA gene from *F. prausnitzii* and from closely related Ruminococcaceae were recovered from GenBank (see Table S1, Supplemental Digital Content 1) and aligned using Clustal W (38). Both primers and hydrolysis probes were manually designed, from consensus sequences (see Table S1, Supplemental Digital Content 1) specifically built for each purpose, after the guidelines set by Applied Biosystems (Foster

City, CA, USA) for the design of primers and probes for allelic discrimination, and further checked using the software Primer Express version 3.0 (Applied Biosystems, Foster City, CA, USA). Oligonucleotides were also evaluated using NetPrimer software (PREMIER Biosoft International, Palo Alto, CA, USA) to check for primer-dimer structures, hairpins and possible cross dimer interactions. Resulting primers and probes are listed in Table 2.

Optimization of the reagents for the *F. prausnitzii* phylogroup qPCR assay was performed as described in Supplementary text (Supplemental Digital Content 1). Oligonucleotides specificity was checked against the Ribosomal Database Project II (RDP) (39) and GenBank database through Seqmatch and BLAST (40), respectively. Coverages were evaluated using the SILVA Probe Match and Evaluation Tool - TestProbe 3.0 (available at http://www.arb-silva.de/search/testprobe/). Finally, *in vitro* inclusivity/exclusivity test was performed including 89 bacterial strains, 9 of which were *F. prausnitzii* (see Table S2, Supplemental Digital Content 1). Linearity, efficiency and detection limit of the assay were determined as detailed in Supplementary text (Supplemental Digital Content 1).

Quantification standards for qPCR

Standard DNA templates from *F. prausnitzii* strain S3L/3 (phylogroup I), *F. prausnitzii* DSM 17677 (phylogroup II) and *E. coli* CECT 105 were prepared as genetic constructs after PCR amplification as previously reported (41, 42), and subsequent insertion of the whole 16S rRNA gene into a pCR4-TOPO cloning plasmid (Invitrogen, Carlsbad, CA, USA) after manufacturer's guidelines. After purification with the NucleoSpin Plasmid (Macherey-Nagel GmbH&Co., Duren, Germany), plasmids were linearized with *SpeI* (*F. prausnitzii*) or *PstI* (*E. coli*) and quantified using Qubit Quantitation Platform (Invitrogen). Initial target concentration was inferred as previously reported (18). Standard curves were obtained from 10-fold serial dilutions of the titrated suspension of linearized plasmids, and ranged from 20 to 2×10^8 copies per reaction, which correspond to the linear dynamic range span for all the reactions (see Supplementary text, Supplemental Digital Content 1). The standard curve built with *F. prausn*itzii DSM 17677 16S rRNA gene was used for both the total bacteria and the total Faecalibacteria 16S rRNA gene quantification, and standard curves obtained from either phylogroup were intercalibrated using the total *F. prausnitzii* primers and probe set. Total bacteria 16S rRNA gene quantification and the *F. prausnitzii* standard curve were used to check the *E. coli* standard curve quantification in order to make sure that results obtained with both standard curves were comparable. Finally, to demonstrate that the new assay correctly quantifies the appropriate ratios of

Table 2. 16S rRNA-targeted primers and probes used in this study.

Target	Primer and Probes [a]			PCR conditions [c]		
	Name	Sequence 5'-3'	Reference	Total cycles	Denaturing (ºC; s)	Annealing and extension (ºC; s)
Bacteria	F_Bact 1369	CGGTGAATACGTTCCCGG	(44)	50	95; 15	60; 60
	R_Prok_1492	TACGGCTACCTTGTTACGACTT				
	P_TM_1389F	**6FAM-CTTGTACACACCGCCCGTC-TAMRA**				
F. prausnitzii (total)	Fpra 428 F	TGTAAACTCCTGTTGTTGAGGAAGATAA	(18)	40	95; 15	60; 60
	Fpra 583 R	GCGCTCCCTTTACACCCA				
	Fpra 493 PR	**6FAM-CAAGGAAGTGACGGCTAACTACGTGCCAG-TAMRA**				
DNA IAC [b]	IAC F	TACGGATGAGGAGGACAAAGGA	(18)	40	95; 15	60; 60
	IAC R	CACTTCGCTCTGATCCATTGG				
	IAC PR	**VIC®-CGCCGCTATGGGCATCGCA-TAMRA**				
E. coli	E.coli 395 F	CATGCCGCGTGTATGAAGAA	(43)	40	95; 15	60; 60
	E.coli 490 R	CGGGTAACGTCAATGAGCAAA				
	E.coli 437 PR	**6FAM-TATTAACTTTACTCCCTTCCTCCCCGCTGAA-TAMRA**				
F. prausnitzii (phylogroups)	Fpra 136F	CTCAAAGAGGGGGACAACAGTT	this study	50	95; 15	64; 60
	Fpra 232R	GCCATCTCAAAGCGGATTG				
	PHG1 180PR	**6FAM-TAAGCCCACGACCCGGCATCG-BHQ1**				
	PHG2 180PR	**JOE-TAAGCCCACRGCTCGGCATC-BHQ1**				

[a] Probe sequences are in bold. FAM™ (6-carboxyfluorescin), VIC® (6-carboxyrhodamine), JOE (4',5'-dichloro-2',7'-dimethoxy-5(6)-carboxyfluorescein), TAMRA™ (tetramethylrhodamin) BHQ1 (Black Hole Quencher1).

[b] IAC, Internal Amplification Control; DNA IAC sequence: 5'-TACGGATGAGGAGGACAAAGGACGCCGCTATGGGCATCGCACCAATGGATCAGAGCGAAGTG-3' (according to Ref 18.).

[c] For all quantitative PCR, an initial step at 50ºC during 2 min was performed for amperase treatment. Also an initial denaturation step was set at 95ºC for 10 min. In quantitative PCR, annealing and extension steps were performed simultaneously.

phylogroups I and II when DNA from both is present in the same sample (as would be expected *in vivo*), mixtures of both DNA templates were prepared (*i.e.* phylogroup I: phylogroup II mixed at 100:0, 25:75, 50:50, 75:25, and 0:100) and quantified as unknown samples. Less than 10% of variation was obtained between the experimental qPCR quantification results with that of the expected quantity (see Table S3, Supplemental Digital Content 1).

qPCR assays

Previously reported 16S rDNA-targeting primers and probe were used for total *F. prausnitzii* (18), *E. coli* (43) and total bacteria (44) quantifications, and amplification reactions were performed as described elsewhere (18, 44, 45). The novel assay for *F. prausnitzii* phylogroups quantification was performed in a total volume of 20 µl reactions containing: 1× TaqMan Universal PCR Master Mix 2× (Applied Biosystems, Foster City, CA, USA), 900 nM of each primer, 300 nM of each probe, and up to 50 ng of genomic DNA template. All primers and probes used in this study and also PCR conditions are detailed in Table 2. Total *F. prausnitzii*, *E. coli,* and total bacteria primers and hydrolysis probes were purchased from Applied Biosystems, whereas primers and hydrolysis probes for *F. prausnitzii* phylogroups were acquired from Biomers (Ulm, Germany). The DNA of the internal amplification control (IAC) was synthesized by Bonsai technologies group (Alcobendas, Spain).

Samples were run in duplicate in the same plate. For data analysis, the mean of the duplicate quantifications was used. Duplicates were considered valid if the standard deviation between quantification cycles (C_q) was <0.34 (*i.e.* a difference of <10% of the quantity was tolerated). Quantification controls consisting of at least 5 reactions with a known number of target genes were performed to assess interrun reproducibility. Inhibition was controlled on total *F. prausnitzii* quantification by adding 10^3 copies of IAC template to each reaction. It was considered that there was no inhibition if the obtained C_q was <0.34 different from those obtained when quantifying the IAC alone for any of the replicates. A no-template control consisting of a reaction without *F. prausnitzii* DNA and a nonamplification control which did not contain any DNA template (either bacterial or IAC) was also included in each run. Negative controls resulted in undetectable C_q values in all cases.

All quantitative PCR were performed using a 7500 Real Time PCR system (Applied Biosystems). Data were collected and analyzed using the 7500 SDS system software version 1.4 (Applied Biosystems). All quantifications were performed under average PCR efficiencies of 89.51±7.06%.

Data normalization and statistical analyses

As regards to qualitative analyses, absence of *F. prausnitzii* or its phylogroups were considered if no detection was obtained during the qPCR analysis, corresponding to samples that carried *F. prausnitzii* or the phylogroups below the detection limit (*i. e.* 106.6, 1.10 and 2.39 16S rRNA genes per reaction for total *F. prausnitzii*, phylogroup I and phylogroup II, respectively). Pearson's χ^2 test was used to compare the prevalence of *F. prausnitzii* and its phylogroups between groups of patients and by IBD disease location.

Referring to quantitative analyses, total *F. prausnitzii,* phylogroups and *E. coli* copy numbers were normalized to the total bacteria 16S rRNA gene copies. Data are given as the $\log_{10}$ of the ratio between 16S rRNA gene copies of the target microorganism and million of total bacterial 16S rRNA genes detected in the same sample.

The nonparametric Kruskal-Wallis test was used to test differences in variables with more than 2 categories such as diagnostics, CD and UC disease location, and current medication. Pairwise comparisons of subcategories of these variables were analyzed using a Mann-Whitney U test. This test was also used to compare, within a subgroup of patients, variables with 2 categories such as activity (active CD and patients with UC patients when Crohn's disease activity index (CDAI)>150 (46) and a Mayo score >3 (47), respectively), and intestinal resection. In addition, the receiver operating characteristic (ROC) curve analysis, a plot of the true positive rate (sensitivity) versus false positive rate (1-specificity), was applied to establish the usefulness of *F. prausnitzii*, and each phylogroup to distinguish amongst different intestinal disorders. The accuracy of discrimination was measured by the area under the ROC curve (AUC). An AUC approaching 1 indicates that the test is highly sensitive and highly specific whereas an AUC approaching 0.5 indicates that the test is neither sensitive nor specific.

Spearman correlation coefficient and significance between the phylogroups quantities and between phylogroups quantities and *E. coli* were calculated. The same statistical method was used to analyze the correlation between each one of the phylogroups with respect to total Faecalibacteria quantity and clinical data such as time (in years) since disease onset.

All the statistical analyses were performed using the SPSS 15.0 statistical package (LEAD Technologies, Inc. Charlotte, NC, USA). Significance levels were established for P values ≤ 0.05.

Ethical consideration

This work was approved by the Ethics Committee of Clinical Research of the Hospital Universitari Dr. Josep Trueta (Girona, Spain) and the Institut d'Assistència Sanitària of Girona (Salt, Spain) on 24th February 2009 and 21st April 2009, respectively. Informed consent from the subjects was obtained before enrollment.

Results

Features of the novel multiplex qPCR assay for *F. prausnitzii* phylogroups I and II

In this study, a novel oligonucleotide set was designed to quantify the 2 recently described *F. prausnitzii* phylogroups (Table 2, and see Supplemental Digital Content 1). The *in silico* analysis of the oligonucleotide set of choice showed that primer Fpra 136F-Fpra 232R were specific for *F. prausnitzii* and targeted all the isolates available to date, whereas the probes PHG1 180PR and PHG2 180PR specifically matched phylogroups I and II, respectively. These results were confirmed *in vitro* by the inclusivity-exclusivity tests (see Table S2, Supplemental Digital Content 1). Coverage of the Fpra 136F-Fpra 232R primers set was 74.85% of the sequences in the SILVA data sets. PHG1 180PR probe targeted 20.50% of the *Faecalibacterium* sp. sequences whereas PHG2 180PR probe coverage was 38.80% of the *Faecalibacterium* sp. sequences in this database. For both reactions reliable quantification was possible over a linear range span of 7 logarithms, starting at 20 target genes per reaction (R^2=0.998), with an average efficiency of 85.68±3.23 % for phylogroup I and 90.31±3.40% for the phylogroup II. The detection limits were 1.10 and 2.39 target genes for phylogroup I and phylogroup II, respectively.

Prevalence of mucosa-associated *F. prausnitzii* and phylogroups I and II along the gut in health and disease

Prevalence of *F. prausnitzii* and both phylogroups as calculated from positive determinations over total samples was analyzed both by disease status considering all data across all sites (Figure 1), and by sample location (Figure 2). *F. prausnitzii* prevalence was lower in patients with CD than in H (Figure 1). Patients with CD and I-CD feature lower *F. prausnitzii* prevalence than those with E1, E2, E3 and colonic CD (C-CD). Prevalence values ranged from 81% to 100%, except for I-CD whose value was significantly lower (down to 68%, P≤0.046) regardless of the location (Figure 2). In contrast, reduced prevalence was only evident in ileal and colonic samples in ileocolonic-CD (IC-CD) (75% and 80% respectively) and in rectal samples in C-CD (75%), although the differences were not statistically supported (Figure 2).

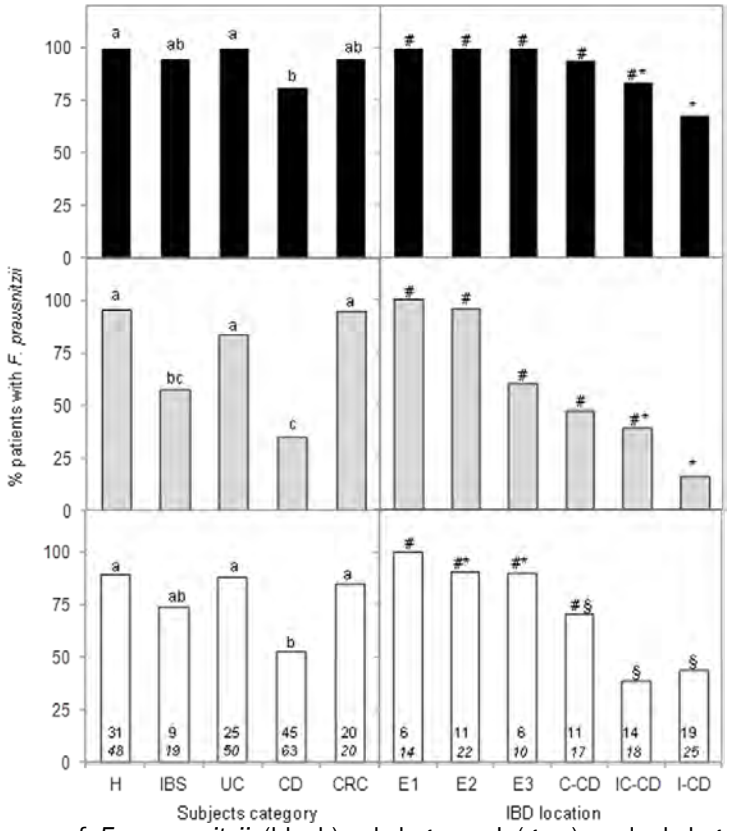


Figure 1. Prevalence of *F. prausnitzii* (black), phylogroup I (gray) and phylogroup II (white) by disease (left) and inflammatory bowel disease location (right) considering all bieopsies from different gut locations. H, control subjects; CRC, colorectal cancer; IBS, irritable bowel syndrome; UC, ulcerative colitis; CD, Crohn's disease; E1, proctitis; E2, left-sided colitis; E3, pancolitis; C-CD, colonic CD; IC-CD, ileocolonic CD; I-CD, ileal CD, IBD, inflammatory bowel disease. Numbers in the bars indicate the number of patients (*biopsies*) analyzed to calculate the prevalence. Statistics was calculated separately for each panel. Homogeneous subgroups (P>0.05) within each panel are indicated with the same symbols above the bars, whereas groups of patients with statistically different prevalence (P<0.05) do not share any superscript.

As far as the phylogroups are concerned, both were found to be less prevalent in patients with CD (P<0.001) than in the H and CRC groups, particularly in those with ileal involvement (Figure 1). Of particular interest is the absence of phylogroup I from all 5 ileal samples of the patients with I-CD analyzed (Figure 2). Phylogroup II was less prevalent in patients with I-CD regardless of sample location. The same happened in colon and ileum of patients with IC-CD, and also in rectal samples of patients with C-CD. For CRC and patients with UC, the prevalence remained similar to H. Nevertheless, phylogroup I showed a trend of lower values in ulcerative pancolitis, which did not reach statistical significance (P=0.053) probably because of the low number of samples processed. Similarly patients with IBS only had reduced prevalence of phylogroup I in comparison with H subjects.

Both phylogroups cooccurred in 85.4% and 85.0% of samples containing *F. prausnitzii* from H and patients with CRC, respectively. Phylogroup I was exclusive in 10% of H and CRC subjects, whereas phylogroup II was found as the only representative

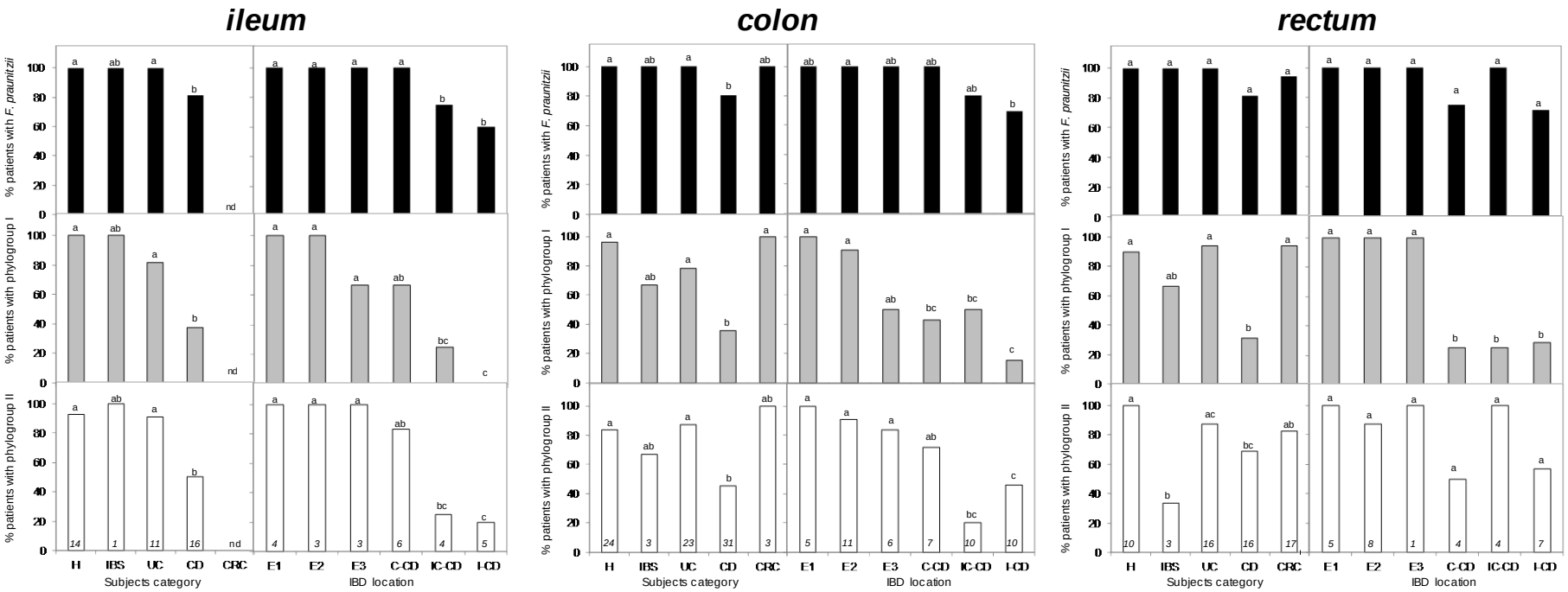


Figure 2. *F. prausnitzii* (black), phylogroup I (gray) and phylogroup II (white) prevalence at ileum, colon, and rectum by disease (left panels) and inflammatory bowel disease location (right panels). H, control subjects; CRC, colorectal cancer; IBS, irritable bowel syndrome; UC, ulcerative colitis; CD, Crohn's disease; E1, proctitis; E2, left-sided colitis; E3, pancolitis; C-CD, colonic CD; IC-CD, ileocolonic CD; I-CD, ileal CD, IBD, inflammatory bowel disease; nd, not determined. Numbers in the bars indicate the number of biopsies analyzed to calculate the prevalence. Homogeneous subgroups (P > 0.05) within each panel are indicated with the same symbol above the bars, whereas groups of patients with statistically different prevalences (P<0.05) do not share any superscript.

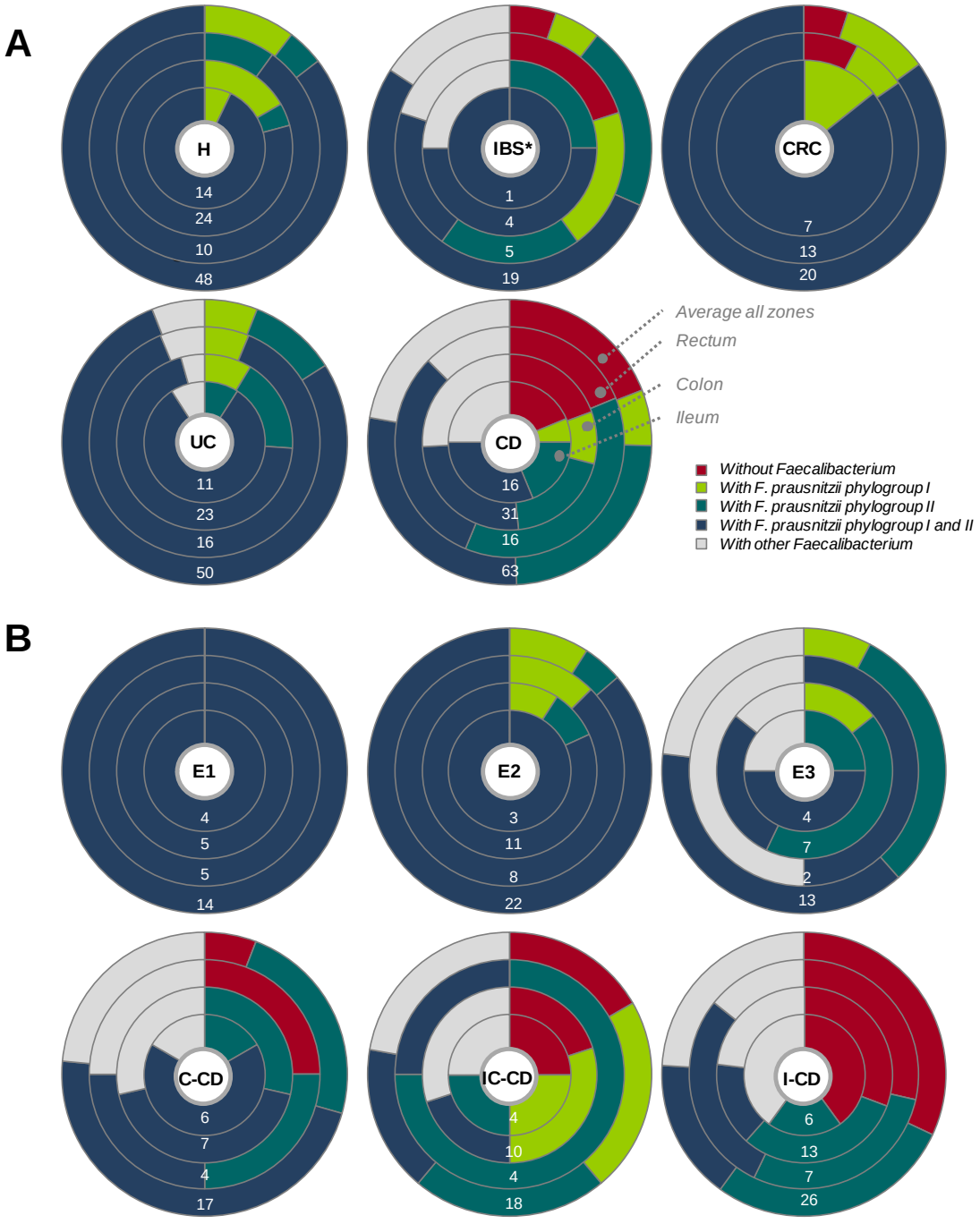


Figure 3. Prevalence of *F. prausnitzii*, phylogroup I, and phylogroup II in each group of patients (A) and by disease locations in patients with and inflammatory bowel disease (IBD). (B). Prevalences along the gut (from inner to outer circles-ileum, colon and rectum) and pooling all the samples (outer circle) have been represented. H, control subjects; CRC, colorectal cancer; IBS, irritable bowel syndrome; UC, ulcerative colitis; CD, Crohn's disease; E1, proctitis; E2, left-sided colitis; E3, pancolitis; C-CD, colonic CD; IC-CD, ileocolonic CD; I-CD, ileal CD. Numbers in the sectors indicate the number of biopsies analyzed.
* Samples with uncertain location have been included in the average analysis of IBS patients.

in 4.2% of H subjects (Figure 3A). In contrast, 16% of patients with IBS, &% of patients with UC, and 22% of patients with CD and *F. prausnitzii* carried neither phylogroup I nor II, which suggests the existence of other phylogroups. Differences in prevalences were

observed between IBD disease locations. All patients with less severe UC (*i.e.* E1 and E2) had one or both *F. prausnitzii* phylogroups, resembling H subjects, whereas none of the phylogroups were detected in 23.1% of ulcerative pancolitis patients despite having *F. prausnitzii* (Figure 3B). Similarly, 22.2% of all patients with CD did not show either of the phylogroups. Within patients with CD, 47.1% of patients with C-CD had both *F. prausnitzii* phylogroups whereas the presence of a unique phylogroup was more frequent (44.4% of patients with IC-CD and 28.0% of patients with I-CD) in those with ileal involvement. Remarkably whenever a single phylogroup was found in I-CD, it always was the phylogroup II.

Abundances of mucosa-associated *F. prausnitzii* and phylogroups in health and disease

The abundance of *F. prausnitzii* and its phylogroups from all the biopsies pooled together was compared amongst patients with different intestinal disorders and H subjects (Table 3). *F. prausnitzii* was less abundant in patients with IBD and CRC as compared with healthy subjects (P<0.001), whereas patients with IBS closely resembled the H group. As previously reported (18), within patients with UC, those with E1 and E3 presented *F. prausnitzii* loads similar to H subjects, whereas those with E2 had abundances between patients with CD and H subjects. In patients with CD, those with ileal involvement presented the lowest levels of this bacterium, whereas patients with C-CD similar to UC (Table 3).

Table 3. Abundances of mucosa-associated *F. prausnitzii* and its phylogroups in controls (H), and patients with Irritable Bowel Syndrome (IBS), Ulcerative Colitis (UC), and Crohn's disease (CD).

	n patients *(n biopsies)*	*F. prausnitzii**§	Phylogroup I*§	Phylogroup II*§
H	**31 *(48)***	**5.33±0.58 [a]**	**3.39±0.87 [a]**	**3.39±1.51 [a]**
IBS	**9 *(19)***	**5.29±0.54 [a,b]**	**2.53±1.22 [a,b]**	**2.72±1.06 [a,b]**
CRC	**20 *(20)***	**4.42±0.58 [c]**	**2.66±0.91 [b]**	**2.56±1.14 [a,b]**
UC	**25 *(50)***	**5.00±0.62 [b]**	**2.59±1.24 [b]**	**2.93±0.99 [a]**
Location				
Ulcerative proctitis (E1)	6 *(14)*	5.09±0.29 [a]	2.76±0.38 [a,b]	3.22±0.43 [a]
Distal UC (E2)	11 *(22)*	4.49±0.59 [b]	2.58±1.15 [a, b]	2.84±0.93 [a,b]
Extensive UC (E3)	6 *(10)*	5.34±0.69 [a]	0.95±1.60 [b,c]	3.13±1.02 [a,b]
CD	**45 *(63)***	**4.26±1.34 [c]**	**0.71±1.65 [c]**	**1.54±1.47 [c]**
Location				
Ileal-CD (L1)	19 *(25)*	3.97±1.42 [c]	0.43±1.33 [c]	1.14±1.54 [b]
Colonic-CD (L2)	11 *(17)*	5.06±1.07 [a,c]	1.54±1.71 [b c]	2.63±1.51 [a,b]
Ileocolonic-CD (L3)	14 *(18)*	4.30±1.12 [b, c]	1.06±1.72 [b,c]	1.38±1.54 [b]

Disease locations of UC and CD patients are analyzed as independent groups.
* Statistics was calculated separately for each variable (column). Groups of patients with similar abundances of *F. prausnitzii* or its phylogroups are indicated with the same superscript (a,b or c) whereas groups not sharing superscript are those with statistically different median abundance values (P<0.05)
§ Median $\log_{10}$ 16S rRNA gene copies/ million bacterial 16S rRNA gene copies ± standard deviations

F. prausnitzii phylogroup I load was reduced in all the intestinal diseases analyzed in comparison to H subjects, except for patients with IBS, probably because of the low number of patients included and the high dispersion of data. This reduction was particularly conspicuous in patients with CD, who had values 1000 times lower than H subjects ($P<0.001$). When analyzing data by disease location, all patients with CD showed this marked reduction of phylogroup I abundance, and also those patients with UC and E3 who resembled more to patients with CD than to those with other UC disease location. In contrast, *F. prausnitzii* phylogroup II abundance was only significantly reduced in patients with CD in comparison to H ($P<0.001$) (Table 3), particularly in those with ileal involvement (either I-CD or IC-CD), suggesting that in these patients, the depletion of *F. prausnitzii* affects the overall Faecalibacteria community.

Interestingly, in H subject, and patients with CRC and IBS, the abundance of the 2 phylogroups was similar, whereas in patients with IBD, phylogroup II outnumbered phylogroup I (Table 3). In patients with UC, *F. prausnitzii* phylogroup II abundance was twice that of phylogroup I, whereas in patients with CD, the imbalance between the 2 phylogroups was more marked, and with *F. prausnitzii* phylogroup II, the abundance was 6.76 times higher than that of phylogroup I. Notably, patients with E3 also featured a marked imbalance in phylogroup abundances, which resembled that found in CD.

Usefulness of mucosa-associated *F. prausnitzii* and phylogroup abundance as diagnostic biomarkers

ROC curve analysis, applied to test the putative accuracy of total *F. prausnitzii* abundance as an indicator to differentiate between 2 groups of patients, confirmed that the reduction of this species load is a good discriminator for patients with CRC from H and patients with IBS, with AUC values greater than 0.8 (Figure 4) and with an 80% of specificity and above 70% of sensitivity at a set threshold. Good discrimination was also observed between patients with CD and H, although for the same specificity values, sensitivity was reduced to 62%. Interestingly, phylogroup I abundance was a more accurate indicator to distinguish H from subjects with IBD, than total *F. prausnitzii* abundance (Figure 4). When comparing H subjects with UC more than 76.60% of sensitivity and above 57.14% of specificity at a set threshold were reached for all the disease locations but with the exception of ulcerative proctitis (E1). Specificity was improved up to 70% when considering exclusively E3 patients. In addition, phylogroup I abundance was a particularly accurate biomarker to distinguish H and patients with CD (91.48% sensitivity, and 73.02% specificity), especially those with I-CD in which 91.48% sensitivity and up to 88.00% of specificity could be reached. Although phylogroup II abundance can accurately discriminate H and subjects with CD, AUC values were slightly lower than those obtained for phylogroup I, thus indicating

that the latter is a more suitable biomarker for H status. In contrast, phylogroup II was a useful biomarker to discriminate within IBD subtypes because the best AUC values were obtained to distinguish between ulcerative pancolitis patients and those with CD and colonic involvement (phylogroup II AUC E3vsC-CD=0.817).

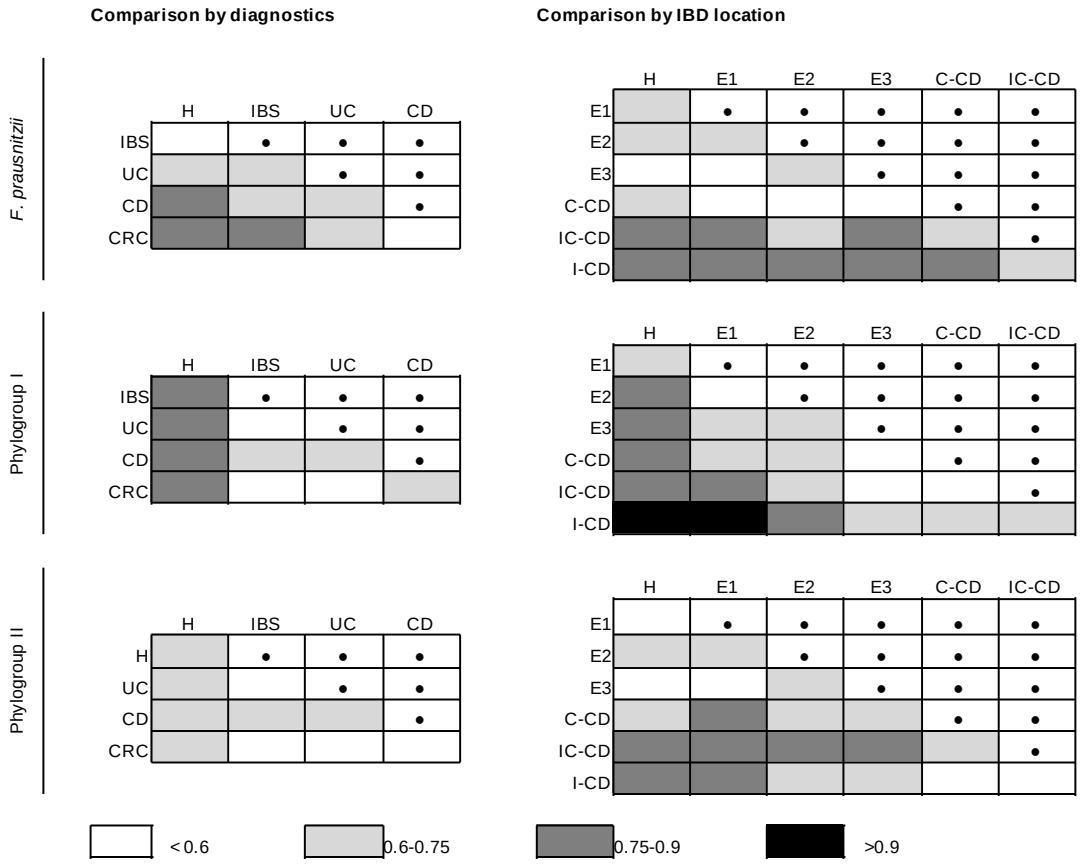


Figure 4. Suitability of *F. prausnitzii* and phylogroups abundances as biomarkers to distinguish amongst different intestinal disorders and inflammatory bowel disease locations determined by the area under the curve (AUC) obtained by receiver operating characteristic analysis (ROC curve). A test is considered to be suitable if the AUC ranges from 0.6 to 0.75, and to have good sensitivity and specificity if the AUC range from 0.75 to 0.9. (H, controls; IBD, inflammatory bowel disease; IBS, irritable bowel syndrome; UC, ulcerative colitis; CD, Crohn's disease; CRC, colorectal cancer; I-CD, ileal CD; IC-CD, ileocolonic CD, C-CD, colonic CD; E1, ulcerative proctitis, E2, distal UC; E3, extensive UC or pancolitis).

Correlation of total *F. prausnitzii* with phylogroups, between phylogroups, and between phylogroups and *E. coli* abundances

Correlations between total *F. prausnitzii* and phylogroups abundances were conducted to determine wether the depletion in *F. prausnitzii* abundance could be attributed to the depletion of one of the phylogroups in certain intestinal disorders. In H subjects and patients with IBD, a positive correlation exists between the 2 phylogroups and total *F. prausnitzii* abundance, suggesting that they are key contributors to *F. prausnitzii* abundance in the gut of these groups of patients (Table 4). In contrast, in patients with CRC, a

significant correlation was found only for phylogroup I and total Faecalibacteria abundance, which suggests that phylogroup II subpopulation is not particularly influencing total *F. prausnitzii* load in this clinical condition. Similarly, no significant correlation was found in IBS, probably because of the low cohort of these patients included in this study.

Abundances of the 2 phylogroups were positively correlated in H and patients with IBD. In contrast, no significant correlation between the 2 phylogroup loads was found in patients with IBS and CRC (Table 4), suggesting that in these disorders, the gut environmental conditions differentially impact on each phylogroup.

Table 4. Correlation between *F. prausnitzii* and its phylogroups abundances, and between phylogroups abundances in controls (H), and patients with Irritable Bowel Syndrome (IBS), Ulcerative Colitis (UC), and Crohn's disease (CD).

Diagnostics	N patients (*N biopsies*)	*F. prausnitzii* vs phylogroup I		*F. prausnitzii* vs phylogroup II		Phylogroup I vs phylogroup II	
		ρ	P	ρ	P	ρ	P
H	31 *(48)*	0.573	<0.001	0.741	<0.001	0.716	<0.001
CRC	20 *(20)*	0.626	0.003	0.177	0.456	0.190	0.422
IBS	9 *(19)*	0.327	0.172	0.284	0.239	0.217	0.373
UC	25 *(50)*	0.671	<0.001	0.677	<0.001	0.667	<0.001
CD	45 *(63)*	0.618	<0.001	0.743	<0.001	0.565	<0.001

Finally, correlation between *F. prausnitzii* phylogroups and *E. coli* was determined to establish wether they were positively or negatively correlated, and whether this could provide supporting evidence about a putative common factor affecting negatively/positively both bacterial populations in a given patient or about a direct/indirect effect of one population over the other. No statistically significant correlation between *E. coli* and any of the two *F. prausnitzii* phylogroups was found in any group of patients or by IBD location. However, it is intriguing that phylogroup II load negatively correlated with *E. coli* in all the groups of gut disease (Table S4, Supplemental Digital Content 1). A significant negative correlation between phylogroup II and *E. coli* across all disease groups together was observed (ρ=-0.196, P=0.016).

F. prausnitzii and phylogroup abundances in relation to patients clinical and treatment data

F. prausnitzii and the abundance of the phylogroups did not differ between active and inactive patients with UC (see Table S5, Supplemental Digital Content 1). Although no statistical significance was reached, active patients with CD showed a marked reduction on phylogroup I abundance with respect to patients with CD in remission (P=0.106).

F. prausnitzii abundance was reduced in those patients with CD that underwent intestinal resection (see Table S6, Supplemental Digital Content 1). Interestingly, this could be attributable to lower phylogroup II abundance, which was 10-fold lower in resected

patients with CD than in those without intestinal surgery (P=0.001), whereas the phylogroup I load was only slightly lower between resected and non-resected patients.

Concerning disease duration, no statistically significant correlation was found between time from disease onset and *F. prausnitzii* and phylogroup abundances (data not shown).

Finally, as far as therapies are concerned, data were analyzed taking into account the medication of the patients at the time of sampling (see Table S7, Supplemental Digital Content 1). No differences in *F. prausnitzii* or in phylogroup abundances were observed between medications within any IBD. However, those patients with CD who received no treatment or mesalazine had higher *F. prausnitzii* loads than those patients under moderate immunosupressants or anti-tumor necrosis factor. No medication was associated with the recovery of normal levels of these bacterial indicators.

Discussion

In this study we have analyzed the prevalence and abundance of mucosa associated *F. prausnitzii* and its 2 phylogroups in H, and subjects with IBS, CRC and IBD, taking into account both the diversity of disease locations and the clinical features of patients. *F. prausnitzii* abundance is reduced in several intestinal disorders, and for the first time, we describe how the abundance of its phylogroups differs between intestinal conditions, and in relation to *E. coli*. New data on phylogroup distribution along the gut and in relation to clinical data are revealed.

Our data show that mucosa-associated *F. prausnitzii* loads are markedly reduced in CRC and patients with CD, especially in those with ileal involvement. *F. prausnitzii* was below detection limits of the method (106.6 16S rRNA genes of *F. prausnitzii* per reaction) in 5% of patients with CRC and 20% of patients with CD. Patients with UC also featured a lower *F. prausnitzii* abundance than H subjects, but this depletion was 4-times less prominent than the depletion observed in patients with CD and CRC. Finally, abundance in patients with IBS was similar to H subjects. Our study is in agreement with previous reports, which found *F. prausnitzii* to be less abundant and/or prevalent in adult CD (11, 17-21, 28, 34), UC (7, 19, 24-27) and CRC (33). Intriguingly, a recent study has reported increased *F. prausnitzii* abundance in *de-novo* pediatric patients with CD (22), which is not in line with our results and suggests that dysbiosis in adult and pediatric CD may be different, which merits further investigation. Contradictory data can also be found in the literature concerning lower *F. prausnitzii* numbers in CRC (33, 48, 49). Controversy also exists with respect to patients with IBS. Some previous studies report normal counts (7, 20, 23, 50-53), whereas others found lower numbers, but exclusively in those patients with IBS of alternating type (30). We have

not observed depletion in *F. prausnitzii* load in patients with IBS, although this observation could be biased by the small cohort size, which also had not been classified by disease type.

Among many intestinal disorders (IBS, diarrhea, upper gut disorder, colonic disorder, UC, CD, ischemic colitis, celiac disease and self-limiting colitis), patients with CD have been shown to possess the lowest abundance of *F. prausnitzii* in feces (20, 23). These results are now corroborated in intestinal mucosa by our study, which reveals for the first time that at mucosa level, the abundance of *F. prausnitzii* in CRC is similar to that found in patients with CD. Altogether, these findings suggest that down-shifts in *F. prausnitzii* numbers occur under several pathological disorders although the numbers are especially compromised in severe diseases such as CRC and CD. Our study supports the view of *F. prausnitzii* as an indicator of healthy gut status. It has been reported that *F. prausnitzii* is seriously affected by the changes that occur in gut environmental conditions during disease such as changes in pH, bile salt or oxygen (13, 14). This suggests that its decrease may be regarded as an indicator of an altered gut environment, which can be associated with worse disease prognostics, and that changes in the abundance of this species are not necessarily indicating a pathogenic role but rather that yet some environmental factors of the gut compromising its presence remain altered. Besides, a recent study (54) has suggested that the beneficial effect of enteral nutrition in pediatric CD is not mediated by an increase in this species. The fact that mucosal healing can be achieved in CD with enteral nutrition while *F. prausnitzii* decreases suggests that the effect of this species may be relatively modest compared with some other factor(s) that are improved by enteral nutrition.

We have further analyzed the prevalence and abundance of *F. prausnitzii* phylogroups I and II, by developing a new quantitative assay. Approximately 25% of all *Faecalibacterium* sequences available in SILVA data set are not targeted *in silico* by any of these assays. This discrepancy could be due to the existence of other phylogroups and/or because different phylogroup probes do not include all members within each phylogroup. Our results are still valid however to compare between diseases in our study, as all have been analyzed with the same technique. Most H subjects H and with CRC harbored both phylogroups far higher than the detectable level whereas patients with IBS and IBD feature a reduced prevalence of one of the phylogroups, particularly those with CD. Furthermore, phylogroup I and II were undetected in 16% of patients with IBS and 22% of patients with CD and *F. prausnitzii*. These results suggest an imbalance within the *F. prausnitzii* population in these diseases and suggest the existence of at least one more phylogroup.

Quantitative analysis demonstrated that, although the depletion in phylogroup I abundance is a general feature in abnormal gut conditions, the depletion of *F. prausnitzii* phylogroup II seems to be specific to patients with CD with ileal disease location. At this

stage, we cannot determine whether or not this is involved in the pathogenesis of this disease location, or whether it is a consequence. It does however indicate that the overall *Faecalibacteria* community is depleted in patients with CD and supports the hypothesis that patients with ileal disease location constitute a differentiated pathological entity (21). Previous work based on inferring *F. prausnitzii* subgroup quantities from PCR band intensity on agarose gels already suggested that the levels of M21/2 subgroup (phylogroup I) in patients with CD were lower than those in the control group and that the levels of the A2-165 subgroup (phylogroup II) were the lowest for patients with CD (23). These observations have now been quantitatively confirmed on mucosal samples by our study, which, in addition, reveals differences between IBD subtypes. Currently, there is no phenotypic trait that consistently distinguishes *F. prausnitzii* members from one or other phylogroup (14), which can undoubtedly explain their differential load in specific disease phenotypes, although the effect of host factors differentially influencing *F. prausnitzii* subpopulations has not yet been explored. Another hypothesis could be that *F. prausnitzii* phylogroups interact in a different manner with other members of the microbiome. We have observed that in all patients with gut disease phylogroup II tends to negatively correlate with *E. coli,* whereas correlation between this species and phylogroup I depends on the patient group. Our data does not allow us to decipher whether or not one population is directly influencing the other, but suggests that interaction between these 2 species varies between different gut conditions.

The potential use of *F. prausnitzii* and its phylogroup quantification to assist in IBD diagnostics or to monitor disease progression is of interest in clinical management. It has been reported that CD and UC could be differentiated through monitoring *F. prausnitzii* abundance in conjunction with fecal leukocyte counts (20). Furthermore, the usefulness of *F. prausnitzii* abundance in biopsy samples as a biomarker to distinguish patients with IBD from IBS and H subjects has been demonstrated recently (18). Adding *E. coli* counts as a complementary contrasting indicator improved the discrimination power and allowed for good differentiation of IBD locations that are difficult to discriminate, such as I-CD from IC-CD, and C-CD from extensive UC. *F. prausnitzii* phylogroups I and II could be novel biomarkers to improve differential diagnosis of those IBD subtypes which are usually difficult to distinguish. For instance, we have observed that phylogroup II is reduced in IC-CD with respect to C-CD, whereas phylogroup I is less abundant in extensive UC than in distal UC. Moreover, phylogroup I proved to be a more accurate marker than total *F. prausnitzii* counts to discriminate between H subjects and those with IBD. However, prospective studies to support the applicability of *F. prausnitzii* phylogroup abundance as biomarkers by comparing with, for example, established measures such as C- reactive protein, albumin, and fecal calprotectin would be necessary to truly determine their ability to distinguish between intestinal disorders and IBD subtypes. In addition, further validation of

our results in feces would provide a noninvasive approach to identify CD and UC, which is more likely to be used as diagnostics test.

The fact that *F. prausnitzii* abundance, including both phylogroups, seems to remain lower under remission suggests that this depletion may be occurring at early disease stages or even before disease onset and remains altered over time even if there is endoscopic and clinical remission. Previous studies based on biopsies from patients with CD with both active and in remission carry lower *F. prausnitzii* numbers in comparison with H subjects (18, 21). Our data confirm that this feature is shared by both phylogroups. However, despite no statistically significant differences being observed, active patients with CD presented a reduction of phylogroup I levels in comparison with inactive patients. Therefore, subsequent studies on larger cohorts of patients are needed to corroborate this trend, and follow-up studies would also be interesting to determine how disease status may be specifically compromising this subpopulation and to irrefutably rule out its potential usefulness as a prognostic biomarker. In agreement with previous studies (11, 18) lower numbers of *F. prausnitzii* were detected in resected patients with CD. This reduction is also replicated with phylogroups counts. In this case, nevertheless, statistical significant differences were only achieved for phylogroup II, probably because the depletion is more striking. However, whether this shift is a consequence of the surgery is still unclear.

In general terms, we have observed that current medication does not restore the levels of mucosa-associated *F. prausnitzii* or its phylogroups, which is in agreement with a previous report (18) although little attention has been paid in the literature to the effect of medication on *F. prausnitzii* abundance. Some specific therapies not included in this study such as chemotherapy with somatostatin and interferon α-2b treatment in patients with mid-gut neuroendocrine tumor (55), and rifaximin (56), high-dose cortisol and infliximab (20) in patients with CD have proven useful to restore the level of this species. Altogether, these data suggest that such therapies will be more useful in terms of counterbalancing *F. prausnitzii* depletion; follow-up studies monitoring this species load in patients starting these treatments will be necessary to demonstrate their effect on modulating this species and its phylogroups abundance.

Finally, in patients with IBD, *F. prausnitzii* abundance correlated positively with both phylogroup I and II. A positive correlation was also found between phylogroups. This indicates that environmental changes in the gut ecosystem of patients with IBD have a similar effect on both phylogroups and that a reduction in both phylogroups is an indication of the total *F. prausnitzii* population decrease. In line with this observation, all *F. prausnitzii* representatives cultured so far, regardless of their phylogroup, are sensitive to small physico-chemical changes in the gut occurring as a consequence of disease status, such as lower pH

or increased bile salts content (14). However, the depletion of phylogroup II was specifically observed in patients with CD and ileal involvement. This suggests that specific phenomena in particular gut diseases can compromise one group more than the other. Therefore the use of *Faecalibacterium* as a fine indicator of different gut environmental alterations, which would be characteristic of each intestinal disease, could be the subject for further research. In addition, assessing whether or not *F. prausnitzii* populations hosted by patients with different intestinal disorders are different from those found in H subjects at the level of subtype composition may shed light on the role of this species in gut health maintenance.

Conclusion

Mucosa-associated *F. prausnitzii* is significantly depleted in patients with gut disorders. Populations of phylogroups I and II of this species however depend on the disease condition. Thus, while *F. prausnitzii* phylogroup I is generally depleted in most intestinal diseases, phylogoup II numbers are specifically reduced in CD. Phylogroup loads can be potentially applied to assist in gut disease diagnostics and in IBD location classification.

Acknowledgements

We appreciate the generosity of the patients who freely gave their time and samples to make this study possible, and the theatre staff of all centers for their dedication and careful sample collection. The authors thank Ms. Teresa Mas-de-Xaxars for assistance in CRC samples recruitment and Ms. Natàlia Adell from the Serveis Tècnics de Recerca of the Universitat de Girona for statistical assistance.

Conflicts of interest and sources of funding

X. Aldeguer is consultant from AbbVie and has received honoraria for lectures including services on speakers bureaus from AbbVie and MSD. The remaining authors have no conflict of interest to disclose.

Supported in part by the Spanish Ministry of Education and Science through project SAF2010-15896. M. Lopez-Siles was recipient of an FI grant from the Generalitat de Catalunya (2010FI_B2 00135), which receives support from the European Union Commissionate. H. J. Flint and S. H. Duncan acknowledge support from the Scottish Government Food, Land and People program.

References

1. Arumugam M, Raes J, Pelletier E, et al. Enterotypes of the human gut microbiome. Nature. 2011;473:174-180
2. Eckburg PB, Bik EM, Bernstein CN, et al. Diversity of the human intestinal microbial flora. Science. 2005;308:1635-1638

3. Hold GL, Schwiertz A, Aminov RI, et al. Oligonucleotide probes that detect quantitatively significant groups of butyrate-producing bacteria in human feces. Appl Environ Microbiol. 2003;69:4320-4324
4. Schwiertz A, Jacobi M, Frick JS, et al. Microbiota in pediatric inflammatory bowel disease. J Pediatr. 2010;157:240-244
5. Suau A, Rochet V, Sghir A, et al. *Fusobacterium prausnitzii* and related species represent a dominant group within the human fecal flora. Systematic and Applied Microbiology. 2001;24:139-145
6. Walker AW, Ince J, Duncan SH, et al. Dominant and diet-responsive groups of bacteria within the human colonic microbiota. ISME J. 2011; 5 (2):220-230
7. Swidsinski A, Loening-Baucke V, Lochs H, et al. Spatial organization of bacterial flora in normal and inflamed intestine: a fluorescence in situ hybridization study in mice. World J Gastroenterol. 2005;11:1131-1140
8. Nava GM, Stappenbeck TS. Diversity of the autochthonous colonic microbiota. Gut Microbes. 2011;2:99-104
9. Baumgart M, Dogan B, Rishniw M, et al. Culture independent analysis of ileal mucosa reveals a selective increase in invasive *Escherichia coli* of novel phylogeny relative to depletion of Clostridiales in Crohn's disease involving the ileum. ISME J. 2007;1:403-418
10. Louis P, Flint HJ. Diversity, metabolism and microbial ecology of butyrate-producing bacteria from the human large intestine. FEMS Microbiol Lett. 2009;294:1-8
11. Sokol H, Pigneur B, Watterlot L, et al. *Faecalibacterium prausnitzii* is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. Proc Natl Acad Sci U S A. 2008;105:16731-16736
12. Barcenilla A, Pryde SE, Martin JC, et al. Phylogenetic relationships of butyrate-producing bacteria from the human gut. Appl Environ Microbiol. 2000;66:1654-1661
13. Duncan SH, Hold GL, Harmsen HJ, et al. Growth requirements and fermentation products of *Fusobacterium prausnitzii*, and a proposal to reclassify it as *Faecalibacterium prausnitzii* gen. nov., comb. nov. Int J Syst Evol Microbiol 2002;52:2141-2146
14. Lopez-Siles M, Khan TM, Duncan SH, et al. Cultured representatives of two major phylogroups of human colonic *Faecalibacterium prausnitzii* can utilize pectin, uronic acids, and host-derived substrates for growth. Appl Environ Microbiol. 2012;78:420-428
15. Carlsson AH, Yakymenko O, Olivier I, et al. *Faecalibacterium prausnitzii* supernatant improves intestinal barrier function in mice DSS colitis. Scand J Gastroenterol. 2013;48:1136-1144
16. Wrzosek L, Miquel S, Noordine ML, et al. *Bacteroides thetaiotaomicron* and *Faecalibacterium prausnitzii* influence the production of mucus glycans and the development of goblet cells in the colonic epithelium of a gnotobiotic model rodent. BMC Biol. 2013;11:61
17. Miquel S, Martin R, Rossi O, et al. *Faecalibacterium prausnitzii* and human intestinal health. Curr Opin Microbiol. 2013;16:255-261
18. Lopez-Siles M, Martinez-Medina M, Busquets D, et al. Mucosa-associated *Faecalibacterium prausnitzii* and *Escherichia coli* co-abundance can distinguish Irritable Bowel Syndrome and Inflammatory Bowel Disease phenotypes. International Journal of Medical Microbiology. 2014;304:464-475
19. Sokol H, Seksik P, Furet JP, et al. Low counts of *Faecalibacterium prausnitzii* in colitis microbiota. Inflamm Bowel Dis. 2009;15:1183-1189
20. Swidsinski A, Loening-Baucke V, Vaneechoutte M, et al. Active Crohn's disease and ulcerative colitis can be specifically diagnosed and monitored based on the biostructure of the fecal flora. Inflamm Bowel Dis. 2008;14:147-161
21. Willing B, Halfvarson J, Dicksved J, et al. Twin studies reveal specific imbalances in the mucosa-associated microbiota of patients with ileal Crohn's disease. Inflamm Bowel Dis. 2009;15:653-660
22. Hansen R, Russell RK, Reiff C, et al. Microbiota of de-novo pediatric IBD: increased *Faecalibacterium prausnitzii* and reduced bacterial diversity in Crohn's but not in ulcerative colitis. Am J Gastroenterol. 2012;107:1913-1922
23. Jia W, Whitehead RN, Griffiths L, et al. Is the abundance of *Faecalibacterium prausnitzii* relevant to Crohn's disease? FEMS Microbiol Lett. 2010;310:138-144
24. Kabeerdoss J, Sankaran V, Pugazhendhi S, et al. *Clostridium leptum* group bacteria abundance and diversity in the fecal microbiota of patients with inflammatory bowel disease: a case-control study in India. BMC Gastroenterol. 2013;13:20
25. Machiels K, Joossens M, Sabino J, et al. A decrease of the butyrate-producing species *Roseburia hominis* and *Faecalibacterium prausnitzii* defines dysbiosis in patients with ulcerative colitis. Gut. 2014; 63(8).1275-1283
26. McLaughlin SD, Clark SK, Tekkis PP, et al. The bacterial pathogenesis and treatment of pouchitis. Therap Adv Gastroenterol. 2010;3:335-348
27. Vermeiren J, Van den Abbeele P, Laukens D, et al. Decreased colonization of fecal *Clostridium coccoides/Eubacterium rectale* species from ulcerative colitis patients in an in vitro dynamic gut model with mucin environment. FEMS Microbiol Ecol. 2012;79:685-696

28. Frank DN, St Amand AL, Feldman RA, et al. Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases. Proc Natl Acad Sci U S A. 2007;104:13780-13785
29. Nagalingam NA, Lynch SV. Role of the microbiota in inflammatory bowel diseases. Inflamm Bowel Dis. 2012;18:968-984
30. Rajilic-Stojanovic M, Biagi E, Heilig HG, et al. Global and deep molecular analysis of microbiota signatures in fecal samples from patients with irritable bowel syndrome. Gastroenterology. 2011;141:1792-1801
31. Ghoshal UC, Shukla R, Ghoshal U, et al. The gut microbiota and irritable bowel syndrome: friend or foe? Int J Inflam. 2012;2012:151085
32. Spiller RC. Inflammation as a basis for functional GI disorders. Best Practice & Research Clinical Gastroenterology. 2004;18:641-661
33. Balamurugan R, Rajendiran E, George S, et al. Real-time polymerase chain reaction quantification of specific butyrate-producing bacteria, *Desulfovibrio* and *Enterococcus faecalis* in the feces of patients with colorectal cancer. J Gastroenterol Hepatol. 2008;23:1298-1303
34. Martinez-Medina M, Aldeguer X, Gonzalez-Huix F, et al. Abnormal microbiota composition in the ileocolonic mucosa of Crohn's disease patients as revealed by polymerase chain reaction-denaturing gradient gel electrophoresis. Inflamm Bowel Dis. 2006;12:1136-1145
35. Mondot S, Kang S, Furet JP, et al. Highlighting new phylogenetic specificities of Crohn's disease microbiota. Inflamm Bowel Dis. 2011;17:185-192
36. Seksik P, Rigottier-Gois L, Gramet G, et al. Alterations of the dominant faecal bacterial groups in patients with Crohn's disease of the colon. Gut. 2003;52:237-242
37. Silverberg MS, Satsangi J, Ahmad T, et al. Toward an integrated clinical, molecular and serological classification of inflammatory bowel disease: Report of a Working Party of the 2005 Montreal World Congress of Gastroenterology. Can J Gastroenterol. 2005;19 Suppl A:5-36
38. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 1994;22:4673-4680
39. Maidak BL, Cole JR, Lilburn TG, et al. The RDP-II (Ribosomal Database Project). Nucleic Acids Research. 2001;29:173-174
40. Altschul SF, Madden TL, Schäffer AA, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Research. 1997;25:3389-3402
41. Lane DJ. 16S/23S rRNA sequencing. New York: In E. Stackebrandt and M. Goodfellow (ed.), Nucleic acid techniques in bacterial systematics. John Willy and Sons; 1991
42. Weisburg WG, Barns SM, Pelletier DA, et al. 16S ribosomal DNA amplification for phylogenetic study. J Bacteriol. 1991;173:697-703
43. Huijsdens XW, Linskens RK, Mak M, et al. Quantification of Bacteria Adherent to Gastrointestinal Mucosa by Real-Time PCR. J Clin Microbiol. 2002;40:4423-4427
44. Furet J-P, Firmesse O, Gourmelon M, et al. Comparative assessment of human and farm animal faecal microbiota using real-time quantitative PCR. FEMS Microbiology Ecology. 2009;68:351-362
45. Martinez-Medina M, Aldeguer X, Lopez-Siles M, et al. Molecular diversity of Escherichia coli in the human gut: new ecological evidence supporting the role of adherent-invasive E. coli (AIEC) in Crohn's disease. Inflamm Bowel Dis. 2009;15:872-882
46. Best WR, Becktel JM, Singleton JW, et al. Development of a Crohn's disease activity index. National Cooperative Crohn's Disease Study. Gastroenterology. 1976;70:439-444
47. Pineton de Chambrun G, Peyrin-Biroulet L, Lemann M, et al. Clinical implications of mucosal healing for the management of IBD. Nat Rev Gastroenterol Hepatol. 2010;7:15-29
48. Wang T, Cai G, Qiu Y, et al. Structural segregation of gut microbiota between colorectal cancer patients and healthy volunteers. ISME J. 2012;6:320-329
49. Sobhani I, Tap J, Roudot-Thoraval F, et al. Microbial dysbiosis in colorectal cancer (CRC) patients. PLoS One. 2011;6:e16393
50. Duboc H, Rainteau D, Rajca S, et al. Increase in fecal primary bile acids and dysbiosis in patients with diarrhea-predominant irritable bowel syndrome. Neurogastroenterol Motil. 2012;24:513-520, e246-517
51. Kassinen A, Krogius-Kurikka L, Makivuokko H, et al. The fecal microbiota of irritable bowel syndrome patients differs significantly from that of healthy subjects. Gastroenterology. 2007;133:24-33
52. Malinen E, Rinttila T, Kajander K, et al. Analysis of the Fecal Microbiota of Irritable Bowel Syndrome Patients and Healthy Controls with Real-Time PCR. Am J Gastroenterol. 2005;100:373-382
53. Rigsbee L, Agans R, Shankar V, et al. Quantitative Profiling of Gut Microbiota of Children With Diarrhea-Predominant Irritable Bowel Syndrome. Am J Gastroenterol. 2012;107:1740-1751
54. Gerasimidis K, Bertz M, Hanske L, et al. Decline in presumptively protective gut bacterial species and metabolites are paradoxically associated with disease improvement in pediatric Crohn's disease during enteral nutrition. Inflamm Bowel Dis. 2014;20:861-871

55. Dorffel Y, Swidsinski A, Loening-Baucke V, et al. Common biostructure of the colonic microbiota in neuroendocrine tumors and Crohn's disease and the effect of therapy. Inflamm Bowel Dis. 2012;18:1663-1671
56. Maccaferri S, Vitali B, Klinder A, et al. Rifaximin modulates the colonic microbiota of patients with Crohn's disease: an in vitro approach using a continuous culture colonic model system. J Antimicrob Chemother. 2010;65:2556-2565

5 concluding remarks

Current concept of health includes the indigenous microbiota as an essential component [265]. Many studies have shown the connection between gut microbiota and several functions related to maintain host well-being (for review see [59] and references therein). On the other hand, a rising number of studies have evidenced gut microbiota imbalances in those suffering intestinal disorders [27, 37-39, 62, 66-68, 74-88]. However, the complexity of the microbial community found in the gut, its diversity of functions and the fact that the large majority of gut microorganisms have not yet been cultured [19, 266, 267] is preventing to precisely define what constitutes a "healthy microbiome". For this purpose, the study of some of the most abundant species found in the gut, and consistently reported to be altered in gut disorders, such as *F. prausnitzii*, can shed light on this field. In line with these needs, the overall aim of the PhD research described in this Thesis was to obtain novel physiological and ecological insights of the important beneficial gut commensal *F. prausnitzii*, and to further explore its usefulness as biomarker of gut health.

5.1. Factors supporting *F. prausnitzii* presence in the gut.

In the last years there have been a rising number of works reporting *F. prausnitzii* depletion in gut disease [15, 26, 85, 87, 88, 106, 111, 112, 114, 117, 176, 187, 190, 206, 225, 241-244, 268, 269]. On the other hand, many studies have evidenced its key role to maintain gut homeostasis [118, 197, 202, 219, 222, 223, 226]. Therefore, it is of relevance to reveal which gut factors are crucial to support *F. prausnitzii* presence in the gut, and the extent of their influence. The results obtained in this Thesis allow discussing the influence of (i)

carbon sources used for its growth, (ii) effect of gut physicochemical conditions, and (iii) interaction with other gut symbionts.

5.1.1 Carbon sources used by *F. prausnitzii* for growth.

Given the difficulty to culture this species, the number of studies characterising the phenotypical traits of this species has remained scarce. To determine which carbon sources found in the gut can use this species to grow is of relevance to seek out novel strategies to maintain or boost this species population in individuals who feature a depletion of *F. prausnitzii*.

Phenotypic characterisation of 17 *F. prausnitzii* isolates from healthy individuals performed in **Article I** of this work has evidenced that all *F. prausnitzii* tested were able to grow by using simple carbohydrates such as glucose, cellobiose and maltose as carbon source. Some differences exist between strains in their capability to ferment more complex carbohydrates such as those from diet and derived from the host.

Our data show that in general, *F. prausnitzii* strains were able to grow on inulin. Previous studies have evidenced that healthy volunteers feature a 4% increase on this species population after ingestion of 10 g of inulin per day (over a 16-days period) in comparison with a control period without any supplement intake [55]. However, our data suggest selectivity in stimulating some isolates because only two strains (A2-165 and HTF-F, both from phylogroup II) were able to grow well fermenting inulin. Further studies to determine how *F. prausnitzii* population change under this prebiotic intake are required.

F. prausnitzii strains had a limited ability to utilize other polysaccharides that can also be frequently encountered in the gut lumen such as arabinogalactan, xylan and soluble starch [270]. *In vivo* studies on healthy human volunteers employing different prebiotics revealed a clear stimulation of *F. prausnitzii* [48, 54, 55]. It can be therefore hypothesised that *F. prausnitzii* relies in cross-feeding by other members of the gut microbiota.

Interestingly, our data in **Article I** have shown that most of the isolates grew well on apple pectin and are also able to use some pectin derivatives such as galacturonic acid. On one hand this result is supported by information from the reference genome of this species, where pectinolytic enzymes are encoded, precisely a pectin methylesterase and an endopolygalacturonase. On the other hand, an *in vivo* study has shown that Firmicutes are promoted in apple pectin-feed rats [271]. All together suggest that pectin could serve as a potential prebiotic for *F. prausnitzii*. Despite it is well fermented within the human colon, few species of gut microbiota have been reported to have the capability to grow on pectin [272]. *Bacteroides* spp. have been reported to be efficient pectin utilizers [273]. However, our *in vitro*

competition studies including the known pectin utilizers *B. thetaiotaomicron* and *E. eligens* suggest that, under physiological conditions, *F. prausnitzii* can play a key role in fermentation of some types of pectin and that it can compete successfully with other gut bacteria for this substrate.

In addition, *F. prausnitzii* strains were also capable of utilizing the host-derived sugar *N*-acetylglucosamine, and several genes involved in glucosamine utilisation have been found in *F. prausnitzii* S3L/3 genome. Interestingly, it has been reported that treatment with this compound may improve CD because it forms a major part of the glycoproteins incorporated in the mucosal layer of the gastrointestinal tract [274]. Therefore, it will serve as a healing factor in inflamed, damaged soft tissues of the gut. Given the capability to ferment this carbohydrate by *F. prausnitzii*, it would be interesting to explore the effect on restoring beneficial gut bacteria in CD patients undergoing this treatment.

Finally, *F. prausnitzii* isolates were unable to grow *in vitro* on gastric mucin or mucopolysaccharides (heparin, hyaluronic acid, and choindrotin sulphate), suggesting that this species does not contribute to the release of ammonium or sulphate to the gut lumen [275, 276]. Very few bacteria have been reported to be able to use mucin as carbon source (such as *Akkermansia muciniphila* [277-280]) and it is more likely that a consortium of bacteria are required for its full metabolism [281]. Our study does not allow deciphering whether or not *F. prausnitzii* would benefit from mucin metabolism in the gut. Further studies to reveal the interaction of *F. prausnitzii* with mucin-degraders and acetate-producers like *A. muciniphila* would be of interest. Acetate is a required compound for *F. prausnitzii* growth [197], although no data about how meaningful cross-feeding routes of acetat between specific bacteria are, and many bacteria contribute acetate to the overall pool in the gut.

Altogether, our study indicates that *F. prausnitzii* has the ability to switch between diet- and host-derived substrates. The metabolic versatility of this species can explain why this is one of the most abundant species found in the gut [10, 15, 24, 25, 115, 198, 207, 211, 237]. This capability can be explored further to define novel strategies to restore *F. prausnitzii* populations in diseased gut in the future by using some of these carbohydrates alone or combined as prebiotics.

5.1.2 Effect of gut physicochemical conditions

In addition to carbohydrate fermentation profile, we also explored tolerance to changes in gut pH and bile salt concentration as physiological factors that can play a role in determining the ability of an organism to survive in the gut environment. Additionally, these traits might contribute to the temporal/spatial organization of different gut microbes [9].

At low pH values (5.75) the growth of *F. prausnitzii* was generally inhibited, but this phenomenon was found to be strain-dependent (**Article I**). A recent study characterising an extensive collection of *Faecalibacterium* sp. isolates from calves and piglets has also observed that the optimal pH for their growth ranged between 5.5 and 6.7, thus corroborating our findings [264]. On one hand, this suggests that this environmental factor is influencing *F. prausnitzii* distribution along the gut. pH values within the range 5.4-5.8 have been found in the upper parts of the small intestine (duodenum and jejunum) and colon (ascending and transverse) [7, 8]. Therefore, *F. prausnitzii* is more likely to be found in the descending colon and rectum. However, in **Articles III and IV** it has been detected the presence of this species in mucosal samples from terminal ileum. This can be explained because it has been reported that pH from terminal ileum is rather similar to that found in the colon [7, 8]. Besides, other works have reported *F. prausnitzii* in duodenum samples [80]. Therefore, future studies to conclusively determine the presence of this species in upper parts of the gastrointestinal tract are required. On the other hand, it can be also hypothesised that local pH in the gut is compromising the distribution of individual faecalibacterial strains in patients with gut disorders such as IBD, because it has been reported that UC and CD patient often have acidic stools [282, 283].

Similarly, the bile salt tolerance differed among isolates, especially at the lowest concentration tested (0.1%). Inhibition of *Faecalibacterium* sp. growth by the presence of 0.1% of bile salts in the medium has been also observed in the studies of Foditsch *et al.* [264]. In the studies conducted in this Thesis was observed that at this concentration, an average inhibition of 76% in the maximum OD_{650} reached by the cultures, and this value raised up to 97% at 0.5% bile salt concentration. This indicates that in general, *F. prausnitzii* is highly sensitive to a slight increase in physiological concentrations of bile salts, and provides also a plausible explanation for the reduced abundance of faecalibacteria exhibited by CD patients. It has been reported that both bile salts composition and concentrations are altered in CD patients [284, 285]. Further studies to determine if *F. prausnitzii* features higher sensitivity to certain types of bile salts components need to be conducted. Besides, whether or not *F. prausnitzii* has a mechanism to survive at high bile salt concentrations or if it interacts with other bacteria of the gut remains to be elucidated, since other species of intestinal bacteria, such as *Bacteroides* spp. and *Enterococcus faecium* can withstand bile salt concentrations of up to 20% or even 40% respectively [286-288].

Altogether, the findings presented in **Article I** provide a plausible explanation why faecalibacteria exhibit a reduced abundance in CD patients because both, the local pH and bile salt concentrations in the gut, are likely to influence the distribution of individual *F. prausnitzii* strains.

5.1.3. *F. prausnitzii* interaction with some members of gut microbiota

Our results in **Article I** have evidenced that *F. prausnitzii* strains (S3L/3 and A2-165) are able to compete for apple pectin with representatives of the two other known groups of pectin-utilizing bacteria, *B. thetaiotaomicron* and *E. eligens* at three different pHs expected to be found in the large intestine.

In general, the *Bacteroides* spp. are efficient pectin utilizers [273]. In the co-culture experiments performed in this Thesis it was observed that at the highest pH (6.79), where *B. thetaiotaomicron* fermentation of pectin is not curtailed, similar amounts of butyrate to those detected at lower pH values (6.12) were obtained, indicating that *F. prausnitzii* fermentative activity continues despite of the reduced number of *F. prausnitzii* cells counted. It can be hypothesised that initial fermentation of pectin by *B. thetaiotaomicron* can release some pectin derivatives which can then be used by *F. prausnitzii*. In fact, mounting evidence suggest that *F. prausnitzii* might rely in other species like *Bacteroides* for cross-feeding. In line with this, it has been shown that *F. prausnitzii* co-occurs with several members of the *C. coccoides* group and Bacteroidetes in the gut (Figure 5.1).

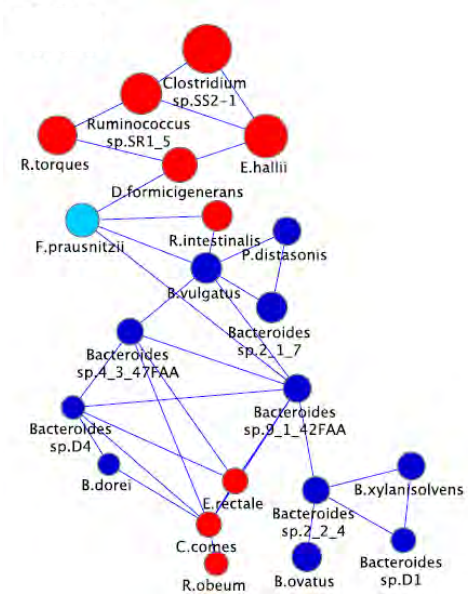


Figure 5.1. Co-occurrence network of *F. prausnitzii* with other human gut bacteria, based on a relative abundance matrix previously reported in Qin *et al.*, 2010 [26]. Co-occurrence modules within the network are defined as a set of species that are connected among each other (directly or via several steps), but not to any other species in the network. Within a module, the nodes represent specie whose genomes have been sequenced. The size of the nodes indicates the average relative abundance across the 124 individuals in the MetaHIT cohort, and the colour of the node reflects taxonomic information (● *Clostridium* Cluster XIVa (*C. coccoides* group) ● *Clostridium* cluster IV (*C. leptum* group) ● Bacteroidetes). Species with significantly positive co-occurrence for any of the six measures used (Pearson, Spearman, Kendall, Bray-Curtis, Euclidean, and mutual information) are joined with an edge.(Adapted from [235]).

Besides, recent studies in mice models have evidenced that *F. prausnitzii* needs the prior presence of *B. thetaiotaomicron* to colonise gut in rat models [223]. The incapability to obtain *F. prausnitzii* mono-associated animal models has been repeatedly observed [289] and very recently a mice model of with *F. prausnitzii* implantation in the gastrointestinal tract following preparation by *E. coli* has been described [228].

Our data in **Article III and IV** also provide some evidence suggesting interaction between *F. prausnitzii* and other members of the gut microbiota such as *E. coli*. While in H

and IBS subjects no correlation between these two species was found, it was intriguing that a correlation between the abundance of these two species exists in IBD patients. In UC patients and those with C-CD *F. prausnitzii* and *E. coli* were positively correlated suggesting that in this condition populations of these two species might be affected by similar host or gut environmental factors. Conversely, a possible negative correlation between *F. prausnitzii* and *E. coli* was observed when analyzing those patients with ileal location of CD (i.e. I-CD and IC-CD). Despite no statistically significant correlation between *E. coli* and any of the two *F. prausnitzii* phylogroups was found in any group of patients or by IBD location, a trend indicating that phylogroup II load negatively correlates with *E. coli* in all the groups of patients with gut disease was observed. This suggested that both species play different roles in gut disease pathogenesis. This hypothesis is sustained by several reports that implicate the AIEC pathotype in CD pathogenesis [122, 123, 290, 291] and those that postulate that a reduction of *F. prausnitzii* might be a crucial factor to enhance disease recurrence [117, 118, 252]. A direct or indirect effect of one population on the other or the effect of changes in gut environment and host factors cannot be ruled out. Therefore, further co-culture experiments engaging the two species would be of assistance to elucidate the interactions between them.

5.2. Taxonomical considerations on *F. prausnitzii*.

5.2.1 *F. prausnitzii* intraspecies diversity

The phylogenetic characterisation of *F. prausnitzii* isolates from healthy individuals performed in **Article I** of this Thesis has evidenced that within this species there exist two phylogroups, which cover 97% of the *F. prausnitzii* 16S rRNA sequences from human faeces that have been recovered through molecular methods [25, 29].

Phenotypical characterisation conducted on *F. prausnitzii* isolates did not reveal any consistent metabolic difference concerning carbohydrate fermentation or tolerance to changes in gut environmental conditions between the members of the two phylogroups (Table 5.1).

However, indications that differences between both groups exist can be found in the literature. First, previous studies indicate a more severe reduction in the phylotypes related to isolate M21/2 (phylogroup I) as compared to phylotypes related to isolate A2-165 (phylogroup II) in biopsies [85] and faecal samples [206] obtained from CD patients. Second, our results have evidenced that differences in prevalence and abundance of both phylogroups exist among patients suffering different gut disorders (**Articles II and IV**). Besides, according to our results in **Article IV** while a depletion in phylogroup I has been observed in active CD patients, lower loads of phylogroup II have been found in CD patients with intestinal resection. This suggests that despite *F. prausnitzii* levels can be

compromised due to several factors, this species might play a role in preventing disease recurrence and acute states of the disease that require surgery as treatment. It remains to be elucidated if both groups play different roles in the disease pathogenesis, and to which extent it is important to have representatives of both.

Table 5.1. Summary of *F. prausnitzii* phylogroups I and II characteristics.

	Phylogroup I	Phylogroup II	p-value
Strains	ATCC27768, M21/2, S3L/3, S4L/4	A2-165, L2-6, L2-15, L2-39, L2-61, HTF-A, HTF-B, HTF-C, HTF-E, HTF-F, HTF-I, HTF-75H, HTF-60C	
Gut distribution	Faces and mucosa	Faces and mucosa	
Genome size (mean bp±SD)	317090±6155	320586±16336	1.000
GC content (% mean±SD)	55.85±0.49	56.45±0.21	0.121
Genes content (mean±SD)	2885.0±97.6	2893.5±103.9	0.439
Proteins content (mean±SD)	2796.5±71.4	2771.0±21.2	1.000
Carbohydrate utilisation (mean OD_{650}±SD)	**n= 4 isolates**	**n= 6 isolates**	
Glucose	0.750±0.311	0.428±0.228	0.163
Cellobiose	0.665±0.277	0.383±0.312	0.240
Maltose	0.685±0.247	0.603±0.273	0.522
Galacturonic acid	0.373±0.208	0.165±0.086	0.110
Galactose	0.435±0.369	0.630±0.183	0.327
Apple pectin	0.408±0.108	0.270±0.224	0.201
Soluble starch	0.075±0.021	0.066±0.011	0.554
Inulin	0.115±0.065	0.510±0.440	0.149
Glucuronic acid	0.150±0.113	0.360±0.410	0.658
N-Acetylgucosamine	0.615±0.224	0.388±0.369	0.221
Glucosamine HCl	0.345±0.177	0.267±0.336	0.134
Tolerance to pH (mean growth rate±SD)			
6.7	0.210±0.070	0.256±.0151	0.755
6.2	0.192±0.050	0.245±0.159	1.000
5.75	0.081±0.039	0.108±0.042	0.110
Tolerance to bile salts (mean maximum OD_{650}±SD)			
0%	0.717±0.427	0.613±0.202	0.977
0.12%	0.174±0.223	0.071±0.150	0.671
0.25%	0.032±0.037	0.014±0.014	0.713
0.5%	0.026±0.033	0.002±0.005	0.089
Metabolites interaction (adapted from [226])	decrease in dihydrothymine and an increase in 4-hydroxyphenylacetylglycine	decreased levels of 3-aminoisobutyrate, taurine, 3,5-hydroxylbenzoate, dimethylamine, 2-hydroxyisobutyrate, glycolate and increased lactate and glycine	
Depletion in gut disorders	In several gut disorders, and in active CD	In CD patients, especially those with intestinal resection.	

Finally, it has been demonstrated that *F. prausnitzii* ATCC2768 (phylogroup I) and *F. prausnitzii* A2-165 (phylogroup II) are related with the modulation of metabolites that influence different host pathways [226]. However, the metabolites associated to each strain

are different (Table 5.1). Although it is not clear if both phylogroups differ in their physiological functions, this study suggests that within *F. prausnitzii* there are members that interact in different manner with the host. Moreover, the link between *F. prausnitzii* and metabolites implicated in the tyrosine metabolism, has been corroborated in faecal samples of healthy subjects by an independent study, and these compounds have been elucidated as metabolic biomarkers able to separate healthy individuals from C-CD patients [227].

The current classification of bacterial species indicates that three criteria should be fulfilled: (i) monophyly, (ii) genomic coherence, and (iii) phenotypic coherence [292]. Our study in **Article I** provides evidences of monophyly but a lack of phenotypic coherence between phylogroups. As concerns to genomic coherence, DNA-DNA hybridization (DDH) has been the gold standard test to assess this criterion. However alternative *in silico* analyses based in full genomes sequencing have been implemented lately such as the average nucleotide identity (ANI). It has been shown that ANI values higher than 94% embraces organisms sharing DDH values higher than 70% which are considered to be genomospecies. For this discussion ANI values between *F. prausnitzii* isolates with sequenced genomes have been calculated (Table 5.2).

Table 5.2. Average nucleotide identity (ANI) values for paired comparisons between *F. prausnitzii* strains whose genome has been fully sequenced. Values corresponding to the same genomospecies are indicated in boldface.

	ANIb* values			
F. prausnitzii isolate	KLE1255	A2-165	L2/6	SL3/3
M21/2	85.26	83.29	82.11	96.70
KLE1255	●	82.79	82.46	84.70
A2-165	82.77	●	82.60	82.74
L2/6	82.33	82.87	●	81.61

	ANIm** values			
F. prausnitzii isolate	KLE1255	A2-165	L2/6	SL3/3
M21/2	89.02	88.52	88.07	97.34
KLE1255	●	88.31	88.65	88.82
A2-165	88.31	●	88.23	88.28
L2/6	88.65	88.23	●	87.99

*ANIb, average nucleotide identity based on BLAST searches of 1 kb genome fragments against a target genome.
** ANIm, average nucleotide identity based on the MUMmer algorithm that does not require the artificial generation of 1kb fragments.
ANIb has better application for distant genomes comparison, while both algorithms give nearly identical values in the high identity range (80-100%).

ANI value between isolates S3L/3 (phylogroup I) and L2/6 (phylogroup II) supports the hypothesis that these would belong to two different genomospecies. Besides, comparison with other draft genomes of *F. prausnitzii* isolates indicate that isolates S3L/3 and M21/2 (both form phylogroup I) share ANI values over 97% confirming that they belong to the same genomospecies. No coherence between A2-165 and L2/6 has been found although the available genome of the former is still pending to be annotated. The accurate sequencing and

annotation of several *F. prausnitzii* strains genomes, as well as assessing some other of these *in silico* parameters such as the digital DDH (dDDH) and the maximal unique matches (MUM) would provide conclusive information in solving whether or not the two phylogroups belong to different genomospecies or genomovars.

5.2.2 Approaching the real diversity of genus *Faecalibacterium*

Studies in **Article I** evidenced that 97.9% of the *F. prausnitzii* sequences recovered by molecular approaches focused on the overall bacterial community in faecal samples [25, 29] belong to either phylogroup I or phylogroup II. This suggested that a 2.1% of these sequences belong to other phylogroups within *F. prausnitzii* or to other species yet to be described within the genus *Faecalibacterium*. Accordingly, our studies in **Article II**, using species-specific primers have evidenced that less that 2% of the sequences recovered do not belong to *F. prausnitzii* phylogroups I or II. These rare phylogroups were mainly recovered from subjects with gut disease. Besides, in **Article IV** it was observed that 16% of IBS, 6% of UC and 22% of CD patients with *F. prausnitzii* carried neither phylogroup I nor phylogroup II, which also suggests the existence of other phylogroups or species within the *Faecalibacterium* genus. Our studies do not allow reaching a consensus about the fraction represented by these other members within *Faecalibacterium* sp.. This may be partially explained by the use of different primers sets between **Articles II and IV**, which may vary slightly in the representatives targeted. Analyses *in silico* against the SILVA database revealed differences in the coverages of *Faecalibacterium* genus sequences between primers sets (coverage Fpra427F-Fpra1127R=70.6%, coverage Fpra428F-Fpra586R=75.7%, coverage Fpra 136F-Fpra232R=74.8%).

Further studies by using next generation sequencing (NGS) would be helpful to corroborate the presence of these rare phylotypes within the faecalibacteria population, and would provide an opportunity to elucidate the taxonomy of the genus *Faecalibacterium*. These analyses should be conducted both, using specific primers for the faecalibacteria population (targeting highly conserved regions of the 16S rRNA gene for the genus, but flanking regions variable enough to identify these rare faecalibacteria members) and using universal bacterial pimers. The former will allow detecting rare members within *Faecalibacterium*. The second is a complementary approach that will offer information on their relative abundance in the bacterial community and will not be biased by known faecalibacteria. Alternatively, 16S rRNA genes within metagenomics sequence datasets, including both healthy volunteers as well as IBD patients (such as that used in study [26]) could be performed as approach to identify further strains related to *F. prausnitzii*. The identification of these novel faecalibacteria members would be of interest because it has been reported that even species with low relative abundance can play pivotal roles in the gut and act as keys-stone species for several processes that take place within the gut [293].

5.3. *F. prausnitzii* populations in healthy and diseased gut

A shared cohort of patients has been engaged in **Articles II-IV**, therefore, results obtained have been discussed together in order to obtain a more complete overview of *F. prausnitzii* populations richness, composition and abundance in gut health and disease.

5.3.1 *F. prausnitzii* population composition and richness

Comparison of *F. prausnitzii* population profile in colonic mucosa of H, IBS, IBD and CRC subjects revealed that the richness of *F. prausnitzii* subtypes was lower in IBD patients than in H subjects. Overall decrease in gut microbiota diversity has been reported previously in the mucosa of IBD patients [97, 106-109, 129]. Besides, Scanlan *et al.* [294] reported a reduced biodiversity in the faecal community of CD patients, mainly attributable to fewer types of Firmicutes detected, and particularly due to a smaller proportion of members from the Clostridial cluster IV (*Clostridium leptum* subgroup). In line with this, in our cohort it was observed that frequently in IBD patients it was detected only one of the two main phylogroups of *F. prausnitzii*. Additionally, in this Thesis has been observed that *F. prausnitzii* populations are not recovered during periods of remission of the disease, suggesting that alterations in this population struggle to normalise with the current patient's treatments. Thereby, new therapies need to be implemented to recover all the diversity of *F. prausnitzii* in these patients.

In **Article II**, it was also revealed that patient groups can be distinguished based on the compositions of *F. prausnitzii* populations and specific phylotypes of each condition were observed. The main members of the *F. prausnitzii* population (four phylotypes, two phylogroups) were detected in all the patients groups, but the distribution of some of them differed between groups of patients. In H and IBS subjects, phylogroup I was more prevalent whereas in IBD and CRC patients it was phylogroup II, and specially the phylotype represented by OTU99_1. These differences in phylotypes and phylogroups prevalence between patients groups allowed us to discriminate patients suffering intestinal disease, especially those with IBD and CRC, from H subjects. Our studies do not allow explaining differences in these phylotypes distributions between groups of patients. In **Article I** it was evidenced that pH and bile salt concentrations influence faecalibacterial strains, but differential effect on the phylogroups of a host factor (yet to be defined) cannot be ruled out. Further studies of isolation and characterisation of strains from patients suffering intestinal disorders would be helpful in order to test *in vitro* whether or not *F. prausnitzii* from different groups of patients have different response to host or gut environmental factors.

5.3.2. *F. prausnitzii* load in healthy and diseased gut

In the last years a growing body of studies has reported *F. prausnitzii* depletion in gut disorders [15, 85, 87, 88, 111, 112, 114, 117, 118, 187, 225, 243, 244]. In this Thesis it has

been analyzed the abundance of mucosa associated *F. prausnitzii* in H, IBS, IBD and CRC subjects, paying careful attention to the diversity of disease phenotypes and clinical features of the patients which has allowed to determine whether or not a depletion in *F. prausnitzii* load can be regarded as a general phenomenon occurring in gut disease.

Results in **Articles III and IV** showed that *F. prausnitzii* loads are markedly reduced in CRC and CD patients, especially in those with ileal involvement. Less prominent depletion in *F. prausnitzii* abundance was observed in UC patients, whereas IBS patients had similar counts to those found in H subjects. Our results are in agreement with the several studies that have reported *F. prausnitzii* depletion in adult CD [85, 87, 111, 114, 117, 118, 225, 246], UC [15, 88, 112, 117, 243, 244] and CRC [187] subjects, and concur with the view that down-shifts in *F. prausnitzii* numbers occur under several pathological disorders. In contrast, some studies where a depletion in *F. prausnitzii* levels in CRC is not observed [187, 190, 295], and some others evidencing increased *F. prausnitzii* abundance in *de-novo* paediatric CD patients [241] have been reported, which is not in line with our results. Besides, a consensus on whether or not IBS patients feature a depletion of *F. prausnitzii* has not been reached since both, studies reporting normal counts [15, 76, 114, 169, 206, 253, 254] and studies reporting lower numbers in IBS patients of alternating type [176] can be found in the literature. Depletion in *F. prausnitzii* load in IBS patients has not been observed in our cohort, but this could be biased by the small cohort size which also had not been classified by disease type. At this stage it cannot be determined the exact role that *F. prausnitzii* plays in the pathogenesis of these diseases. On the one hand an external factor can cause a downshift in *F. prausnitzii*, but also this species depletion can be a contributing fator to disease aggravation. In this case, restoration of normal counts of this species should be explored as a way to achieve healing or attenuate disease progression.

Although the depletion of *F. prausnitzii* is not a specific phenomenon happening in a particular disease, the level of depletion as well as which components of the *F. prausnitzii* population are affected can be different between diseases. In **Article IV** it was demonstrated that, while the depletion in phylogroup I abundance was a general feature in abnormal gut conditions, the depletion of *F. prausnitzii* phylogroup II seems to be specific of CD patients with ileal disease location. Therefore the most plausible hypothesis is that depletion of this species is a phenomenon that happens as consequence of several factors that can affect part or the totality of *F. prausnitzii* members. These differences can be explained for instance if *F. prausnitzii* phylogroups feature different resistance to several diseases, or if they are affected differently by some host factors that may vary between disorders.

Despite it has been reported that *F. prausnitzii* levels recover in faeces during remission [114, 117], the results obtained in this Thesis indicate that total *F. prausnitzii*

depletion in mucosa takes place regardless of activity status of IBD patients which is in line with previous studies [87, 112]. Intriguingly, in active CD patients, *F. prausnitzii* phylogroup I load is specifically compromised. Recently it has been evidenced that Japanese CD patients with high *F. prausnitzii* counts in faeces feature lower CDAI score and CRP levels than those with low counts of this species [246]. In our cohort of CD patients two groups of high *vs.* low *F. prausnitzii* carriers could not be clearly established. Further studies to conclusively determine which clinical data of the patients are improved by the presence of *F. prausnitzii,* and whether or not this is dependent on the quantity of *F. prausnitzii* colonising the gut need to be conducted. Besides, whether or not the presence of this species may be preventing disease chronicity or development towards acute states also remains to be revealed.

5.4. Potential use of *F. prausnitzii* as healthy gut microbiota biomarker.

In **Article III** it was determined the abundance of mucosa-associated *F. prausnitzii* and *E. coli* in different intestinal disorders. Their usefulness in discriminating gut disorders and their correlation with main clinical characteristics of IBD patients was further explored. In **Article IV** a similar study was conducted, but evaluating the usefulness of *F. prausnitzii* phylogroups. For this discussion, results obtained in both studies have been compared in order to reach a consensus about the best biomarker for each condition.

5.4.1. *F. prausnitzii* load as diagnostic supporting tool and IBD subtype biomarker

Total *F. prausnitzii* abundance was the best biomarker to differentiate CRC patients from those with UC and H subjects. In contrast, mucosa-associated *F. prausnitzii* phylogroup I (PHGI) abundance was the best biomarker to discriminate H subjects from those suffering an intestinal disorder (IBS, IBD or CRC) (Table 5.3). Precisely, PHGI abundance was shown to discriminate between H subjects and IBD patients with high accuracy (AUC: 0.816) whereas the suitability of total *F. prausnitzii* abundance (AUC: 0.720) or phylogroup II (PHGII) (AUC: 0.699) was lower.

It is of note that PHGI abundance was a more accurate indicator than total *F. prausnitzii* or PHG II load to distinguish H subjects from patients with CD (Table 5.3), and was shown to be a particularly good indicator of CD with ileal involvement (PHGI AUC: 0.948, Total FP AUC: 0.875 and PHGII AUC: 0.772).

On the other hand, PHGII abundance showed a good discrimination capacity within IBD subtypes. In particular, it was shown to distinguish between ulcerative pancolitis patients (E3) and those with C-CD with suitable accuracy (E3 *vs* C-CD AUC of 0.691), two disorders located in the colon and that may present similar clinical manifestations thus

hampering a clear classification. Due to differences in treatment and management between UC and CD [296] it is of relevance an accurate discrimination between these two entities.

Table 5.3. Usefulness of *F. prausnitzii* (FP) and its phylogroups (PHGI and PHGII) to discriminate between gut disorders and inflammatory bowel disease subtypes.

Discrimination by gut disorders			
	FP	PHGI	PHGII
H vs IBD	0.720	0.816	0.699
H vs IBS+IBD+CRC	0.724	0.804	0.693
HvsIBS	0.591	0.754	0.675
HvsUC	0.681	0.763	0.636
HvsCD	0.750	0.858	0.749
HvsCRC	0.879	0.788	0.674
IBSvsUC	0.601	0.513	0.575
IBSvsCD	0.705	0.723	0.642
IBSvsCRC	0.832	0.505	0.516
UCvsCD	0.646	0.699	0.693
UCvsCRC	0.724	0.505	0.583
CDvsCRC	0.568	0.684	0.591

Discrimination by IBD subtype			
	FP	PHGI	PHGII
E1vsE2	0.656	0.573	0.682
E1vsE3	0.593	0.650	0.504
E1vsC-CD	0.527	0.702	0.748
E1vsIC-CD	0.770	0.790	0.881
E1vsI-CD	0.861	0.917	0.809
E2vsE3	0.709	0.602	0.609
E2vsC-CD	0.561	0.634	0.602
E2vsIC-CD	0.652	0.703	0.761
E2vsI-CD	0.767	0.840	0.713
E3vsC-CD	0.574	0.579	0.691
E3vsIC-CD	0.764	0.572	0.817
E3vsI-CD	0.852	0.732	0.742
C-CDvsIC-CD	0.637	0.513	0.611
C-CDvsI-CD	0.764	0.645	0.589
IC-CDvsI-CD	0.634	0.656	0.530

Controls (H), Irritable Bowel Syndrome (IBS), Ulcerative Colitis (UC), and Crohn's disease (CD) patients, Ulcerative proctitis (E1), Distal UC (E2), Extensive UC or ulcerative pancolitis (E3), Ileal-CD (I-CD), Colonic-CD (C-CD), Ileocolonic-CD (IC-CD).

For IBD patients, and especially those with CD, phylogroup I quantification was the most discriminative biomarker, while for CRC patients total *F. prausnitzii* quantification offered better discrimination from H subjects, therefore, a general biomarker of healthy *vs.* diseased gut status is difficult to be established. Besides phylogroup II load can be a source of additional information to discriminate between IBD subtypes. To validate our observations in a larger cohort of patients, completely independent, that includes volunteers from different ethnicities, would be needed prior to implement *F. prausnitzii* quantification (total or part of this species population) as a tool to assist in gut disease diagnostics. Besides, in future studies it would be interesting to assess if the usefulness of this bacterial indicators to discriminate between disorders is enhanced when used in conjunction with other previously reported biomarkers of intestinal disease such as calprotectin, lactoferrin, C-reactive protein, p-ANCA, and ASCA. To further determine the usefulness of *F. prausnitzii* or its phylogroups as biomarkers to discriminate other intestinal disorders within IBD such as indeterminate colitis, unclassified IBD, pouchitis, microscopic colitis, and/or diverticulosis would be also of interest.

5.4.2. Combination of *F. prausnitzii* with *E. coli* as biomarker.

In **Article III** it was observed that *E. coli* abundance as a complementary contrasting indicator of *F. prausnitzii* abundance improved the discrimination between patients with gut diseases (pair wise comparisons between IBS, UC, CD and CRC) (Table 5.4). Interestingly, it was observed that when comparing particular IBD subtypes, discrimination was improved with the *F. prausnitzii-E. coli* index (F-E index) than when using *F. prausnitzii* alone in all the IBD subtypes comparisons except for those patients with CD of ileal involvement.

Because in **Article IV** it was observed a trend of negative correlation between *F. prausnitzii* phylogroup II and *E. coli*, for this Discussion, it has also been tested whether or not similar indexes including *F. prausnitzii* phylogroups and *E. coli* would enhance discriminating power (Table 5.4).

Table 5.4. Usefulness of *F. prausnitzii* (FP) and its phylogroups (PHGI and PHG II) in conjunction with *E. coli* to discriminate between gut disorders and inflammatory bowel disease subtypes.

Discrimination by gut disorders			
	FP-E	PHGI-E	PHGII-E
HvsIBS	0.580	0.620	0.548
HvsUC	0.614	0.523	0.591
HvsCD	0.690	0.804	0.698
HvsCRC	0.855	0.778	0.741
IBSvsUC	0.552	0.597	0.632
IBSvsCD	0.752	0.731	0.661
IBSvsCRC	0.924	0.687	0.697
UCvsCD	0.769	0.763	0.769
UCvsCRC	0.916	0.740	0.803
CDvsCRC	0.688	0.524	0.559

Discrimination by IBD subtype			
	FP-E	PHGI-E	PHGII-E
E1vsE2	0.513	0.607	0.513
E1vsE3	0.864	0.936	0.771
E1vsC-CD	0.836	0.895	0.933
E1vsIC-CD	0.865	0.849	0.905
E1vsI-CD	0.889	0.940	0.854
E2vsE3	0.727	0.800	0.691
E2vsC-CD	0.777	0.821	0.824
E2vsIC-CD	0.763	0.768	0.816
E2vsI-CD	0.815	0.833	0.775
E3vsC-CD	0.647	0.629	0.712
E3vsIC-CD	0.633	0.567	0.711
E3vsI-CD	0.756	0.628	0.668
C-CDvsIC-CD	0.513	0.552	0.516
C-CDvsI-CD	0.656	0.515	0.520
IC-CDvsI-CD	0.593	0.567	0.516

Controls (H), Irritable Bowel Syndrome (IBS), Ulcerative Colitis (UC), and Crohn's disease (CD) patients, Ulcerative proctitis (E1), Distal UC (E2), Extensive UC or ulcerative pancolitis (E3), Ileal-CD (I-CD), Colonic-CD (C-CD), Ileocolonic-CD (IC-CD).

In general terms, these indexes improved the discrimination of E1 and E2 from all the other IBD subtypes, particularly the PHGI-*E. coli* index. The PHGII-*E. coli* index was shown to discriminate better between E3 and C-CD (AUC: 0.712) than PHGII abundance alone (AUC: 0.691). The inclusion of *E. coli* values did not improve the discrimination achieved for CD of ileal involvement from the other IBD subtypes, probably because discrimination with *F. prausnitzii* load was already accurate.

Finally, *E. coli* can be a promising biomarker of disease recurrence because it was observed higher counts in active than in inactive CD patients, and remission in I-CD patients

was compromised by high abundance of this species. Also our data suggest that this species can be useful as biomarker of treatment efficacy at least in I-CD patients, because it was observed that *E. coli* abundance diminished in I-CD patients that underwent to treatment with anti-TNF α.

The usefulness of gut microbiota assessment to support intestinal diseases diagnostics and or prognostics has gain interest during the last few years. For instance, a study has reported that active CD and UC can be specifically diagnosed monitoring the faecal bacterial community [114], and also a set of six species has been proposed as preliminary diagnostic tool to discriminate active I-CD patients from H [86]. However, to our knowledge no previous study has explored the discrimination between IBD in inactive state, or between subtypes of similar location such as C-CD from UC. In the future, it would be interesting to determine if the combination of *F. prausnitzii* or its phylogroups with other representatives of UC dysbiosis such as *Roseburia hominis* [88] may improve discrimination between IBD subtypes. Alternatively, this can be complemented with the inclusion of other species such as *Ruminococcus gnavus*, *R. torques*, *Dialister invisus* or *Bifidobacterium adolescentis* reported as signature of CD disbiosys [83, 85, 121]. Besides, the usefulness as biomarkers of *F. prausnitzii* quantification in conjunction with other bacteria also indicators of gut health such as *A. muciniphila* merits further investigation. Recently a depletion of this species has also been observed in the mucosa of IBD patients in comparison to control subjects [121], although some other studies have indicated a higher *A. muciniphila* in feces of CD patients [297].

5.4.3. *F. prausnitzii* load as biomarker of disease progression and treatment success.

Our data, although preliminary, allow discussing about the usefulness of *F. prausnitzii* quantification for IBD follow up. First, given the chronic behaviour of IBD, with periods of disease activity and periods of remission, it would be interesting to have a prognostic biomarker for flares-up. Currently, it is unknown if the depletion in *F. prausnitzii* occurs prior to disease development. Several studies agree that depletion of this species occurs in active IBD patients [117, 118, 237, 249, 250], and are in line with our results. In contrast, no consensus for inactive IBD patients has been achieved. Some studies support that this reduction is sustained in remission IBD patients [113, 237] while others have not observed this feature [117, 250]. Our data revealed that, at the mucosal level, inactive IBD patients do not recover *F. prausnitzii* abundance nor phylogroups load to the levels found in H subjects. The fact that *F. prausnitzii* abundance, including both phylogroups, seems to remain lower under remission suggests that this depletion may be occurring at early disease stages or even prior to disease onset, and remains altered over time even if there is endoscopic and clinical remission. It can be hypothesised that differences in the methodology or the cohort engaged

as well as the type of sample analyzed may be confounding factors that are preventing to reach a unanimous result about the usefulness of *F. prausnitzii* to predict flare-ups. Subsequent studies on larger cohorts of patients are needed to corroborate this trend, and follow up studies would also be interesting to determine how disease status may be specifically compromising this subpopulation and to irrefutably rule out its potential usefulness as a prognostic biomarker.

Second, lower numbers of *F. prausnitzii* were detected in resected CD patients in comparison to those without resection, which is in agreement with a previous study [118]. This reduction is also replicated with phylogroups counts. In this case nevertheless, statistical significant differences were only achieved for phylogroup II, probably because the depletion is more notorious. However, whether this shift is a due to the fact that these patients feature a more acute disease, or if it is a consequence of the surgery is still unclear. It would be interesting to conduct follow-up studies to assess the usefulness of this biomarker to precisely predict when such intervention might be needed.

Finally, as far as therapies are concerned treatments with infliximab and high-dose cortisol have been associated with an increase of *F. prausnitzii* levels [114]. In our studies, no medication was associated with the recovery of normal levels of this species suggesting that *F. prausnitzii* would be a poor biomarker to monitor treatment efficacy. However, since our studies are retrospective, further prospective studies are required to establish the usefulness of these biomarkers to monitor long-term treatment efficacy, and to undoubtedly rule out the absence of impact of treatment in this species load in the gut.

5.4.4. Sample of choice and future studies prior to implementation in diagnostics

When analyzing data by sample location, it was observed that colonic biopsies were the most suitable to distinguish disease phenotypes. Although statistical significance was not reached for rectal samples, similar results were obtained. To validate our results in rectal samples would be of interest since rectal sigmoidoscopy is the only non-invasive method currently available to collect tissue samples which will allow implementing mucosa-associated *F. prausnitzii* quantification in routine clinical practice. Alternatively, the validation in samples collected with rectal swabs, which have been reported to have a great similarity to biopsy specimens [14] would also be of interest.

In the future, it would be interesting to determine if faecal *F. prausnitzii* and both phylogroups counts can be also a suitable biomarker for the detection, follow up and/or classification of IBD phenotypes. It has been previously reported that CD and UC could be differentiated through monitoring faecal *F. prausnitzii* abundance in conjunction with leukocyte counts [114]. It should be explored if phylogroup I and/or phylogroup II abundance used in combination with leucocytes counts could also be a useful target and improve diagnostics accuracy.

6 conclusions

From Objective 1

1. *Faecalibacterium prausnitzii* is deeply branched within the *Rumicococcaceae* and includes several phylotypes that are typically represented among the dominant bacteria found in the gut of healthy individuals. The current isolates of this species fall into two phylogroups, well defined phylgenetically, but without consistent phenotypic characteristics that allow their discrimination.

2. *F. prausnitzii* is a nutritionally versatile microorganism because isolates of this species are able to utilise a number of both diet- and host-derived substrates.

3. *F. prausnitzii* can play a vital role in fermentation of some types of pectin in the gut because most of this species isolates grew well on apple pectin, are able to use pectin derivatives such as galacturonic acid, and under physiological conditions, can compete with other pectin utilizers.

4. *F. prausnitzii* isolates are extremely sensitive to small changes in the pH and bile salts concentrations of the colonic environment, which might severely limit the abundance of this beneficial microbe along the gut and in a diseased intestine.

From Objective 2

5. The main phylotypes of *F. prausnitzii* population, that belong to phylogroups I and II, are shared between H and individuals with gut diseases. However, IBD individuals host less rich *F. prausnitzii* populations than H.

6. IBD and CRC *F. prausnitzii* populations can be discriminated from that of H according to the phylotypes composition. This is attributable to increased prevalence of the common phylotype OTU99_1, and to the presence of some disease-specific phylotypes.

7. Imbalance in phylogroups (OTU97_1 and OTU97_2), and abundance of specific phylotypes can be used as biomarkers to distinguish some intestinal diseases as IBD or CRC.

From Objective 3

8. Total *F. prausnitzii* and phylogroup I are depleted in CD, UC and CRC in comparison to H, whilst phylogroup II is specifically reduced in CD. Within IBD, those CD patients with ileal involvement have the lowest *F. prausnitzii* abundances, whereas those with C-CD have values similar to UC patients.

9. Total *F. prausnitzii* abundance was the best biomarker to differentiate CRC patients from those with UC and H subjects. In contrast, mucosa-associated *F. prausnitzii* Phylogroup I (PHGI) abundance was the best biomarker to discriminate H subjects from those suffering an intestinal disorder (either IBS, IBD or CRC). On the other hand, PHGII abundance showed a good discrimination capacity within IBD subtypes.

10. *E. coli* abundance as a complementary contrasting indicator of *F. prausnitzii* abundance improved the discrimination between patients with gut diseases. In general terms when comparing particular IBD subtypes discrimination was improved with the *F. prausnitzii*-*E. coli* index than when using *F. prausnitzii* alone.

11. IBD patients treated with mesalazine and imunosupressants do not display healthy-like levels of *F. prausnitzii* in the mucosa. In contrast *E. coli* levels were reduced in I-CD patients treated with anti-TNF-α. Further prospective studies are required to establish the usefulness of these biomarkers to monitor long-term treatment efficacy.

12. Lower numbers of *F. prausnitzii* were detected in resected CD patients, and this reduction was also replicated with phylogroups counts, specifically for phylogroup II.

Further investigations are required to determine whether this is a cause or a consequence of this intervention.

13. *F. prausnitzii* and the abundance of the phylogroups did not differ between active and inactive UC patients whereas active CD patients showed a marked reduction on phylogroup I abundance with respect to CD patients in remission. Follow up studies would be interesting to determine how disease status may be specifically compromising this subpopulation and to irrefutably rule out its potential usefulness as a prognostic biomarker.

References

This section includes all references cited in the Thesis, except those specifically referenced in Articles I-IV.

1. **Biasucci, G., M. Rubini, S. Riboni,** *et al.*, Mode of delivery affects the bacterial community in the newborn gut. Early Hum Dev, 2010. **86** Suppl 1: p. 13-5.
2. **Dominguez-Bello, M.G., E.K. Costello, M. Contreras,** *et al.*, Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. Proceedings of the National Academy of Sciences, 2010. **107** (26): p. 11971–11975
3. **Heavey, P.M., S.A. Savage, A. Parrett,** *et al.*, Protein-degradation products and bacterial enzyme activities in faeces of breast-fed and formula-fed infants. Br J Nutr, 2003. **89**(4): p. 509-15.
4. **Ringel-Kulka, T., J. Cheng, Y. Ringel,** *et al.*, Intestinal microbiota in healthy U.S. young children and adults--a high throughput microarray analysis. PLoS One, 2013. **8**(5): p. e64315.
5. **Flint, H.J., K.P. Scott, S.H. Duncan,** *et al.*, Microbial degradation of complex carbohydrates in the gut. Gut Microbes, 2012. **3**(4): p. 289-306.
6. **Duncan, S.H. and H.J. Flint**, Probiotics and prebiotics and health in ageing populations. Maturitas, 2013. **75**(1): p. 44-50.
7. **Sekirov, I., S.L. Russell, L.C. Antunes,** *et al.*, Gut microbiota in health and disease. Physiol Rev, 2010. **90**(3): p. 859-904.
8. **Ahmed, S., G.T. Macfarlane, A. Fite,** *et al.*, Mucosa-associated bacterial diversity in relation to human terminal ileum and colonic biopsy samples. Appl Environ Microbiol, 2007. **73**(22): p. 7435-42.
9. **Parfrey, L.W. and R. Knight,** Spatial and temporal variability of the human microbiota. Clin Microbiol Infect, 2012. **18** Suppl 4: p. 8-11.
10. **Eckburg, P.B., E.M. Bik, C.N. Bernstein,** *et al.*, Diversity of the human intestinal microbial flora. Science, 2005. **308**(5728): p. 1635-8.
11. **Green, G.L., J. Brostoff, B. Hudspith,** *et al.*, Molecular characterization of the bacteria adherent to human colorectal mucosa. J Appl Microbiol, 2006. **100**(3): p. 460-9.
12. **Zoetendal, E.G., A. von Wright, T. Vilpponen-Salmela,** *et al.*, Mucosa-associated bacteria in the human gastrointestinal tract are uniformly distributed along the colon and differ from the community recovered from feces. Appl Environ Microbiol, 2002. **68**(7): p. 3401-7.
13. **Wang, M., S. Ahrne, B. Jeppsson,** *et al.*, Comparison of bacterial diversity along the human intestinal tract by direct cloning and sequencing of 16S rRNA genes. FEMS Microbiol Ecol, 2005. **54**(2): p. 219-31.
14. **Albenberg, L., T.V. Esipova, C.P. Judge,** *et al.*, Correlation between intraluminal oxygen gradient and radial partitioning of intestinal microbiota. Gastroenterology, 2014. **147**(5): p. 1055-63 e8.
15. **Swidsinski, A., V. Loening-Baucke, H. Lochs,** *et al.*, Spatial organization of bacterial flora in normal and inflamed intestine: a fluorescence in situ hybridization study in mice. World J Gastroenterol, 2005. **11**(8): p. 1131-40.
16. **Flint, H.J.**, The impact of nutrition on the human microbiome. Nutr Rev, 2012. **70** Suppl 1: p. S10-3.
17. **Flint, H.J., K.P. Scott, P. Louis,** *et al.*, The role of the gut microbiota in nutrition and health. Nat Rev Gastroenterol Hepatol, 2012. **9**(10): p. 577-89.
18. **Rajilic-Stojanovic, M. and W.M. de Vos**, The first 1000 cultured species of the human gastrointestinal microbiota. FEMS Microbiol Rev, 2014. **38**(5): p. 996-1047.
19. **Suau, A., R. Bonnet, M. Sutren,** *et al.*, Direct analysis of genes encoding 16S rRNA from complex communities reveals many novel molecular species within the human gut. Appl. Environ. Microbiol., 1999. **65**(11): p. 4799-807.
20. **Hold, G.L., S.E. Pryde, V.J. Russell,** *et al.*, Assessment of microbial diversity in human colonic samples by 16S rDNA sequence analysis. FEMS Microbiol Ecol, 2002. **39**(1): p. 33-9.
21. **Hayashi, H., M. Sakamoto, and Y. Benno,** Phylogenetic analysis of the human gut microbiota using 16S rDNA clone libraries and strictly anaerobic culture-based methods. Microbiol Immunol, 2002. **46**(8): p. 535-48.
22. **Duncan, S.H., P. Louis, and H.J. Flint**, Cultivable bacterial diversity from the human colon. Lett Appl Microbiol, 2007. **44**(4): p. 343-50.
23. **Rajilic-Stojanovic, M., H. Smidt, and W.M. de Vos**, Diversity of the human gastrointestinal tract microbiota revisited. Environ Microbiol, 2007. **9**(9): p. 2125-36.

24. **Arumugam, M., J. Raes, E. Pelletier,** *et al.*, Enterotypes of the human gut microbiome. Nature, 2011. **473**(7346): p. 174-80.
25. **Walker, A.W., J. Ince, S.H. Duncan,** *et al.*, Dominant and diet-responsive groups of bacteria within the human colonic microbiota. ISME J, **2011**(5): p. 220-230.
26. **Qin, J., R. Li, J. Raes,** *et al.*, A human gut microbial gene catalogue established by metagenomic sequencing. Nature, 2010. **464**(7285): p. 59-65.
27. **Sokol, H. and P. Seksik,** The intestinal microbiota in inflammatory bowel diseases: time to connect with the host. Current Opinion in Gastroenterology, 2010. **26**(4): p. 327-331.
28. **Ley, R.E., C.A. Lozupone, M. Hamady**, *et al.*, Worlds within worlds: evolution of the vertebrate gut microbiota. Nat Rev Microbiol, 2008. **6**(10): p. 776-88.
29. **Tap, J., S. Mondot, F. Levenez,** *et al.*, Towards the human intestinal microbiota phylogenetic core. Environ. Microbiol., 2009. **11**(10): p. 2574-84.
30. **Zoetendal, E.G., A.D. Akkermans, and W.M. De Vos**, Temperature gradient gel electrophoresis analysis of 16S rRNA from human fecal samples reveals stable and host-specific communities of active bacteria. Appl Environ Microbiol, 1998. **64**(10): p. 3854-9.
31. **Ben-Amor, K., H. Heilig, H. Smidt,** *et al.*, Genetic diversity of viable, injured, and dead fecal bacteria assessed by fluorescence-activated cell sorting and 16S rRNA gene analysis. Appl Environ Microbiol, 2005. **71**(8): p. 4679-89.
32. **Vanhoutte, T., G. Huys, E. Brandt,** *et al.*, Temporal stability analysis of the microbiota in human feces by denaturing gradient gel electrophoresis using universal and group-specific 16S rRNA gene primers. FEMS Microbiol Ecol, 2004. **48**(3): p. 437-46.
33. **Wu, G.D., J. Chen, C. Hoffmann,** *et al.*, Linking long-term dietary patterns with gut microbial enterotypes. Science, 2011. **334**(6052): p. 105-8.
34. **Mueller, S., K. Saunier, C. Hanisch,** *et al.*, Differences in fecal microbiota in different European study populations in relation to age, gender, and country: a cross-sectional study. Appl Environ Microbiol, 2006. **72**(2): p. 1027-33.
35. **Ding, T. and P.D. Schloss,** Dynamics and associations of microbial community types across the human body. Nature, 2014. **509**(7500): p. 357-60.
36. **Lozupone, C.A., J.I. Stombaugh, J.I. Gordon**, *et al.*, Diversity, stability and resilience of the human gut microbiota. Nature, 2012. **489**(7415): p. 220-30.
37. **Turnbaugh, P.J., M. Hamady, T. Yatsunenko,** *et al.*, A core gut microbiome in obese and lean twins. Nature, 2009. **457**(7228): p. 480-4.
38. **Ley, R.E., F. Backhed, P. Turnbaugh,** *et al.*, Obesity alters gut microbial ecology. Proc Natl Acad Sci U S A, 2005. **102**(31): p. 11070-5.
39. **Arthur, J.C. and C. Jobin,** The struggle within: microbial influences on colorectal cancer. Inflamm Bowel Dis, 2011. **17**(1): p. 396-409.
40. **Halfvarson, J., T. Jess, A. Magnuson,** *et al.*, Environmental factors in inflammatory bowel disease: a co-twin control study of a Swedish-Danish twin population. Inflamm Bowel Dis, 2006. **12**(10): p. 925-33.
41. **Dicksved, J., J. Halfvarson, M. Rosenquist,** *et al.*, Molecular analysis of the gut microbiota of identical twins with Crohn's disease. ISME J, 2008. **2**(7): p. 716-27.
42. **Nam, Y.D., M.J. Jung, S.W. Roh,** *et al.*, Comparative analysis of Korean human gut microbiota by barcoded pyrosequencing. PLoS One, 2011. **6**(7): p. e22109.
43. **De Filippo, C., D. Cavalieri, M. Di Paola,** *et al.*, Impact of diet in shaping gut microbiota revealed by a comparative study in children from Europe and rural Africa. Proc Natl Acad Sci U S A, 2010. **107**(33): p. 14691-6.
44. **Clemente, J.C., L.K. Ursell, L.W. Parfrey,** *et al.*, The impact of the gut microbiota on human health: an integrative view. Cell, 2012. **148**(6): p. 1258-70.
45. **Rafii, F., J.B. Sutherland, and C.E. Cerniglia,** Effects of treatment with antimicrobial agents on the human colonic microflora. Ther Clin Risk Manag, 2008. **4**(6): p. 1343-58.
46. **Dethlefsen, L., S. Huse, M.L. Sogin,** *et al.*, The pervasive effects of an antibiotic on the human gut microbiota, as revealed by deep 16S rRNA sequencing. PLoS Biol, 2008. **6**(11): p. e280.
47. **Duncan, S.H., A. Belenguer, G. Holtrop**, *et al.*, Reduced dietary intake of carbohydrates by obese subjects results in decreased concentrations of butyrate and butyrate-producing bacteria in feces. Appl Environ Microbiol, 2007. **73**(4): p. 1073-8.
48. **Benus, R.F., T.S. van der Werf, G.W. Welling,** *et al.*, Association between *Faecalibacterium prausnitzii* and dietary fibre in colonic fermentation in healthy human subjects. Br J Nutr, 2010. **104**(5): p. 693-700.
49. **Martinez-Medina, M., J. Denizot, N. Dreux,** *et al.*, Western diet induces dysbiosis with increased *E coli* in CEABAC10 mice, alters host barrier function favouring AIEC colonisation. Gut, 2013. **63**(1): p.116-24.

50. **Wallace, T.C., F. Guarner, K. Madsen,** *et al.*, Human gut microbiota and its relationship to health and disease. Nutr Rev, 2011. **69**(7): p. 392-403.
51. **Abe, A.M., P.J. Gregory, D.J. Hein,** *et al.*, Survey and Systematic Literature Review of Probiotics Stocked in Academic Medical Centers within the United States. Hosp Pharm, 2013. **48**(10): p. 834-47.
52. **Floch, M.H., W.A. Walker, K. Madsen,** *et al.*, Recommendations for probiotic use-2011 update. J Clin Gastroenterol, 2011. **45** Suppl: p. S168-71.
53. **Klaenhammer, T.R., M. Kleerebezem, M.V. Kopp,** *et al.*, The impact of probiotics and prebiotics on the immune system. Nat Rev Immunol, 2012. **12**(10): p. 728-34.
54. **Hooda, S., B.M. Boler, M.C. Serao,** *et al.*, 454 pyrosequencing reveals a shift in fecal microbiota of healthy adult men consuming polydextrose or soluble corn fiber. J Nutr, 2012. **142**(7): p. 1259-65.
55. **Ramirez-Farias, C., K. Slezak, Z. Fuller**, *et al.*, Effect of inulin on the human gut microbiota: stimulation of *Bifidobacterium adolescentis* and *Faecalibacterium prausnitzii*. Br. J. Nutr., 2009. **101**(4): p. 541-50.
56. **Bouhnik, Y., L. Raskine, G. Simoneau,** *et al.*, The capacity of nondigestible carbohydrates to stimulate fecal bifidobacteria in healthy humans: a double-blind, randomized, placebo-controlled, parallel-group, dose-response relation study. Am J Clin Nutr, 2004. **80**(6): p. 1658-64.
57. **Sood, A., V. Midha, G.K. Makharia,** *et al.*, The probiotic preparation, VSL#3 induces remission in patients with mild-to-moderately active ulcerative colitis. Clin Gastroenterol Hepatol, 2009. **7**(11): p. 1202-9.
58. **Sartor, R.B.**, Microbial influences in inflammatory bowel diseases. Gastroenterology, 2008. **134**(2): p. 577-94.
59. **Sears, C.L.**, A dynamic partnership: celebrating our gut flora. Anaerobe, 2005. **11**(5): p. 247-51.
60. **Gill, S.R., M. Pop, R.T. Deboy,** *et al.*, Metagenomic analysis of the human distal gut microbiome. Science, 2006. **312**(5778): p. 1355-9.
61. **Chassard, C., K.P. Scott, P. Marquet,** *et al.*, Assessment of metabolic diversity within the intestinal microbiota from healthy humans using combined molecular and cultural approaches. FEMS Microbiol Ecol, 2008. **66**(3): p. 496-504.
62. **Backhed, F., H. Ding, T. Wang,** *et al.*, The gut microbiota as an environmental factor that regulates fat storage. Proc Natl Acad Sci U S A, 2004. **101**(44): p. 15718-23.
63. **Claus, S.P., S.L. Ellero, B. Berger,** *et al.*, Colonization-induced host-gut microbial metabolic interaction. MBio, 2011. **2**(2): p. e00271-10.
64. **Round, J.L. and S.K. Mazmanian,** The gut microbiota shapes intestinal immune responses during health and disease. Nat Rev Immunol, 2009. **9**(5): p. 313-23.
65. **Maynard, C.L., C.O. Elson, R.D. Hatton,** *et al.*, Reciprocal interactions of the intestinal microbiota and immune system. Nature, 2012. **489**(7415): p. 231-41.
66. **Ley, R.E., P.J. Turnbaugh, S. Klein,** *et al.*, Microbial ecology: human gut microbes associated with obesity. Nature, 2006. **444**(7122): p. 1022-3.
67. **Wen, L., R.E. Ley, P.Y. Volchkov,** *et al.*, Innate immunity and intestinal microbiota in the development of Type 1 diabetes. Nature, 2008. **455**(7216): p. 1109-13.
68. **Larsen, N., F.K. Vogensen, F.W. van den Berg,** *et al.*, Gut microbiota in human adults with type 2 diabetes differs from non-diabetic adults. PLoS One, 2010. **5**(2): p. e9085.
69. **Dumas, M.E., R.H. Barton, A. Toye,** *et al.*, Metabolic profiling reveals a contribution of gut microbiota to fatty liver phenotype in insulin-resistant mice. Proc Natl Acad Sci U S A, 2006. **103**(33): p. 12511-6.
70. **Sidhu, H., M.J. Allison, J.M. Chow,** *et al.*, Rapid reversal of hyperoxaluria in a rat model after probiotic administration of Oxalobacter formigenes. J Urol, 2001. **166**(4): p. 1487-91.
71. **Abdollahi-Roodsaz, S., L.A.B. Joosten, M.I. Koenders,** *et al.*, Stimulation of TLR2 and TLR4 differentially skews the balance of T cells in a mouse model of arthritis. The Journal of Clinical Investigation, 2008. **118**(1): p. 205-216.
72. **Kang, D.W., J.G. Park, Z.E. Ilhan,** *et al.*, Reduced incidence of Prevotella and other fermenters in intestinal microflora of autistic children. PLoS One, 2013. **8**(7): p. e68322.
73. **Song, Y., C. Liu, and S.M. Finegold,** Real-time PCR quantitation of clostridia in feces of autistic children. Appl Environ Microbiol, 2004. **70**(11): p. 6459-65.
74. **Uronis, J.M., M. Muhlbauer, H.H. Herfarth,** *et al.*, Modulation of the intestinal microbiota alters colitis-associated colorectal cancer susceptibility. PLoS One, 2009. **4**(6): p. e6026.
75. **Malinen, E., L. Krogius-Kurikka, A. Lyra,** *et al.*, Association of symptoms with gastrointestinal microbiota in irritable bowel syndrome. World J Gastroenterol, 2010. **16**(36): p. 4532-40.
76. **Malinen, E., T. Rinttila, K. Kajander,** *et al.*, Analysis of the Fecal Microbiota of Irritable Bowel Syndrome Patients and Healthy Controls with Real-Time PCR. Am J Gastroenterol, 2005. **100**(2): p. 373-382.
77. **Parkes, G.C., J. Brostoff, K. Whelan,** *et al.*, Gastrointestinal microbiota in irritable bowel syndrome: their role in its pathogenesis and treatment. Am J Gastroenterol, 2008. **103**(6): p. 1557-67.

78. **Salonen, A., W.M. de Vos, and A. Palva,** Gastrointestinal microbiota in irritable bowel syndrome: present state and perspectives. Microbiology, 2010. **156**(11): p. 3205-3215.
79. **De Palma, G., I. Nadal, M. Medina,** *et al.*, Intestinal dysbiosis and reduced immunoglobulin-coated bacteria associated with coeliac disease in children. BMC Microbiology, 2010. **10**(1): 63.
80. **Nadal, I., E. Donat, C. Ribes-Koninckx,** *et al.*, Imbalance in the composition of the duodenal microbiota of children with coeliac disease. J Med Microbiol, 2007. **56**(Pt 12): p. 1669-74.
81. **Andoh, A., T. Tsujikawa, M. Sasaki,** *et al.*, Faecal microbiota profile of Crohn's disease determined by terminal restriction fragment length polymorphism analysis. Aliment Pharmacol Ther, 2009. **29**(1): p. 75-82.
82. **Gophna, U., K. Sommerfeld, S. Gophna,** *et al.*, Differences between tissue-associated intestinal microfloras of patients with Crohn's disease and ulcerative colitis. J Clin Microbiol, 2006. **44**(11): p. 4136-41.
83. **Joossens, M., G. Huys, M. Cnockaert,** *et al.*, Dysbiosis of the faecal microbiota in patients with Crohn's disease and their unaffected relatives. Gut, 2011. **60**(5): p. 631-7.
84. **Manichanh, C., L. Rigottier-Gois, E. Bonnaud,** *et al.*, Reduced diversity of faecal microbiota in Crohn's disease revealed by a metagenomic approach. Gut, 2006. **55**(2): p. 205-11.
85. **Martinez-Medina, M., X. Aldeguer, F. Gonzalez-Huix,** *et al.*, Abnormal microbiota composition in the ileocolonic mucosa of Crohn's disease patients as revealed by polymerase chain reaction-denaturing gradient gel electrophoresis. Inflamm Bowel Dis, 2006. **12**(12): p. 1136-45.
86. **Mondot, S., S. Kang, J.P. Furet,** *et al.*, Highlighting new phylogenetic specificities of Crohn's disease microbiota. Inflamm Bowel Dis, 2011. **17**(1): p. 185-92.
87. **Willing, B., J. Halfvarson, J. Dicksved,** *et al.*, Twin studies reveal specific imbalances in the mucosa-associated microbiota of patients with ileal Crohn's disease. Inflamm. Bowel. Dis., 2009. **15**(5): p. 653-60.
88. **Machiels, K., M. Joossens, J. Sabino,** *et al.*, A decrease of the butyrate-producing species Roseburia hominis and *Faecalibacterium prausnitzii* defines dysbiosis in patients with ulcerative colitis. Gut, 2013. **63**(8):p.1275-83
89. **Mendoza Hernández JL, L.S.R., Díaz-Rubio M,** *Definiciones y manifestaciones clínicas generales*. Third ed. Enfermedad Inflamatoria Intestinal, ed. F.G. MA Gasull, J Hinojosa, A Obrador. 2007, Madrid: Arán Ediciones S.L.
90. **Wilson, M.**, *Microbial inhabitants of humans*. First ed. 2005, Cambridge: Cambridge University press.
91. **Karlinger, K., T. Gyorke, E. Mako,** *et al.*, The epidemiology and the pathogenesis of inflammatory bowel disease. Eur J Radiol, 2000. **35**(3): p. 154-67.
92. **Silverberg, M.S., J. Satsangi, T. Ahmad,** *et al.*, Toward an integrated clinical, molecular and serological classification of inflammatory bowel disease: Report of a Working Party of the 2005 Montreal World Congress of Gastroenterology. Can J Gastroenterol, 2005. **19** Suppl A: p. 5-36.
93. **Manichanh, C., N. Borruel, F. Casellas,** *et al.*, The gut microbiota in IBD. Nat Rev Gastroenterol Hepatol, 2012. **9**(10): p. 599-608.
94. **Sartor, R.B.**, Mechanisms of disease: pathogenesis of Crohn's disease and ulcerative colitis. Nat Clin Pract Gastroenterol Hepatol, 2006. **3**(7): p. 390-407.
95. **Baumgart, D.C. and W.J. Sandborn**, Crohn's disease. Lancet, 2012. **380**(9853): p. 1590-605.
96. **Boyapati, R., J. Satsangi, and G.T. Ho**, Pathogenesis of Crohn's disease. F1000Prime Rep, 2015. 7: p. 44.
97. **Seksik, P., H. Sokol, P. Lepage,** *et al.*, Review article: the role of bacteria in onset and perpetuation of inflammatory bowel disease. Aliment Pharmacol Ther, 2006. **24** Suppl 3: p. 11-8.
98. **Mizoguchi, A. and E. Mizoguchi**, Inflammatory bowel disease, past, present and future: lessons from animal models. J Gastroenterol, 2008. **43**(1): p. 1-17.
99. **Elson, C.O., Y. Cong, V.J. McCracken,** *et al.*, Experimental models of inflammatory bowel disease reveal innate, adaptive, and regulatory mechanisms of host dialogue with the microbiota. Immunol Rev, 2005. **206**: p. 260-76.
100. **Wirtz, S. and M.F. Neurath,** Mouse models of inflammatory bowel disease. Adv Drug Deliv Rev, 2007. **59**(11): p. 1073-83.
101. **Rutgeerts, P., K. Goboes, M. Peeters,** *et al.*, Effect of faecal stream diversion on recurrence of Crohn's disease in the neoterminal ileum. Lancet, 1991. **338**(8770): p. 771-4.
102. **Sandborn, W.J., C.J. Landers, W.J. Tremaine**, *et al.*, Antineutrophil cytoplasmic antibody correlates with chronic pouchitis after ileal pouch-anal anastomosis. Am J Gastroenterol, 1995. **90**(5): p. 740-7.
103. **Ogura, Y., D.K. Bonen, N. Inohara,** *et al.*, A frameshift mutation in NOD2 associated with susceptibility to Crohn's disease. Nature, 2001. **411**(6837): p. 603-6.
104. **Cadwell, K., J.Y. Liu, S.L. Brown,** *et al.*, A key role for autophagy and the autophagy gene Atg16l1 in mouse and human intestinal Paneth cells. Nature, 2008. **456**(7219): p. 259-63.
105. **Kabi, A., K.P. Nickerson, C.R. Homer,** *et al.*, Digesting the genetics of inflammatory bowel disease: insights from studies of autophagy risk genes. Inflamm Bowel Dis, 2012. **18**(4): p. 782-92.

106. **Sokol, H., C. Lay, P. Seksik,** *et al.*, Analysis of bacterial bowel communities of IBD patients: what has it revealed? Inflamm Bowel Dis, 2008. **14**(6): p. 858-67.
107. **Tamboli, C.P., C. Neut, P. Desreumaux,** *et al.*, Dysbiosis in inflammatory bowel disease. Gut, 2004. **53**(1): p. 1-4.
108. **Chassaing, B. and A. Darfeuille-Michaud**, The commensal microbiota and enteropathogens in the pathogenesis of inflammatory bowel diseases. Gastroenterology, 2011. **140**(6): p. 1720-28.
109. **Barnich, N. and A. Darfeuille-Michaud,** Role of bacteria in the etiopathogenesis of inflammatory bowel disease. World J Gastroenterol, 2007. **13**(42): p. 5571-6.
110. **Bibiloni, R., M. Mangold, K.L. Madsen,** *et al.*, The bacteriology of biopsies differs between newly diagnosed, untreated, Crohn's disease and ulcerative colitis patients. J Med Microbiol, 2006. **55**(8): p. 1141-1149.
111. **Frank, D.N., A.L. St Amand, R.A. Feldman,** *et al.*, Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases. Proc. Natl. Acad. Sci. U S A., 2007. **104**(34): p. 13780-5.
112. **Kabeerdoss, J., V. Sankaran, S. Pugazhendhi,** *et al.*, Clostridium leptum group bacteria abundance and diversity in the fecal microbiota of patients with inflammatory bowel disease: a case-control study in India. BMC Gastroenterol, 2013. **13**: p. 20.
113. **Kang, S., S.E. Denman, M. Morrison,** *et al.*, Dysbiosis of fecal microbiota in Crohn's disease patients as revealed by a custom phylogenetic microarray. Inflamm Bowel Dis, 2010. **16**(12): p. 2034-42.
114. **Swidsinski, A., V. Loening-Baucke, M. Vaneechoutte,** *et al.*, Active Crohn's disease and ulcerative colitis can be specifically diagnosed and monitored based on the biostructure of the fecal flora. Inflamm Bowel Dis, 2008. **14**(2): p. 147-61.
115. **Baumgart, M., B. Dogan, M. Rishniw,** *et al.*, Culture independent analysis of ileal mucosa reveals a selective increase in invasive *Escherichia coli* of novel phylogeny relative to depletion of Clostridiales in Crohn's disease involving the ileum. ISME J, 2007. **1**(5): p. 403-18.
116. **Willing, B.P., J. Dicksved, J. Halfvarson,** *et al.*, A pyrosequencing study in twins shows that gastrointestinal microbial profiles vary with inflammatory bowel disease phenotypes. Gastroenterology, 2010. **139**(6): p. 1844-1854 e1.
117. **Sokol, H., P. Seksik, J.P. Furet,** *et al.*, Low counts of *Faecalibacterium prausnitzii* in colitis microbiota. Inflamm Bowel Dis, 2009. **15**(8): p. 1183-1189.
118. **Sokol, H., B. Pigneur, L. Watterlot,** *et al.*, *Faecalibacterium prausnitzii* is an anti-inflammatory commensal bacterium identified by gut microbiota analysis of Crohn disease patients. Proc. Natl. Acad. Sci. U S A, 2008. **105**(43): p. 16731-16736.
119. **Seksik, P., L. Rigottier-Gois, G. Gramet,** *et al.*, Alterations of the dominant faecal bacterial groups in patients with Crohn's disease of the colon. Gut, 2003. 52(2): p. 237-42.
120. **Fujita, H., Y. Eishi, I. Ishige**, *et al.*, Quantitative analysis of bacterial DNA from *Mycobacteria* spp., *Bacteroides vulgatus,* and *Escherichia coli* in tissue samples from patients with inflammatory bowel diseases. J Gastroenterol, 2002. **37**(7): p. 509-16.
121. **Png, C.W., S.K. Linden, K.S. Gilshenan,** *et al.*, Mucolytic bacteria with increased prevalence in IBD mucosa augment in vitro utilization of mucin by other bacteria. Am J Gastroenterol, 2010. **105**(11): p. 2420-8.
122. **Darfeuille-Michaud, A., J. Boudeau, P. Bulois,** *et al.*, High prevalence of adherent-invasive *Escherichia coli* associated with ileal mucosa in Crohn's disease. Gastroenterology, 2004. **127**(2): p. 412-21.
123. **Martinez-Medina, M., X. Aldeguer, M. Lopez-Siles,** *et al.*, Molecular diversity of *Escherichia coli* in the human gut: new ecological evidence supporting the role of adherent-invasive *E. coli* (AIEC) in Crohn's disease. Inflamm Bowel Dis, 2009. **15**(6): p. 872-82.
124. **Langan, R.C., P.B. Gotsch, M.A. Krafczyk,** *et al.*, Ulcerative colitis: diagnosis and treatment. Am Fam Physician, 2007. **76**(9): p. 1323-30.
125. **Baumgart, D.C. and S.R. Carding**, Inflammatory bowel disease: cause and immunobiology. Lancet, 2007. **369**(9573): p. 1627-40.
126. **Andoh, A., H. Imaeda, T. Aomatsu,** *et al.*, Comparison of the fecal microbiota profiles between ulcerative colitis and Crohn's disease using terminal restriction fragment length polymorphism analysis. J Gastroenterol, 2011. **46**(4): p. 479-86.
127. **Nemoto, H., K. Kataoka, H. Ishikawa,** *et al.*, Reduced Diversity and Imbalance of Fecal Microbiota in Patients with Ulcerative Colitis. Digestive Diseases and Sciences, 2012. **57**(11): p. 2955-2964.
128. **Michail, S., M. Durbin, D. Turner,** *et al.*, Alterations in the gut microbiome of children with severe ulcerative colitis. Inflamm Bowel Dis, 2012. **18**(10): p. 1799-808.
129. **Ott, S.J., S. Plamondon, A. Hart,** *et al.*, Dynamics of the mucosa-associated flora in ulcerative colitis patients during remission and clinical relapse. J Clin Microbiol, 2008. **46**(10): p. 3510-3.

130. **Le Gall, G.n.l., S.O. Noor, K. Ridgway,** *et al.*, Metabolomics of Fecal Extracts Detects Altered Metabolic Activity of Gut Microbiota in Ulcerative Colitis and Irritable Bowel Syndrome. Journal of Proteome Research, 2011. **10**(9): p. 4208-4218.
131. **Swidsinski, A., A. Ladhoff, A. Pernthaler,** *et al.*, Mucosal flora in inflammatory bowel disease. Gastroenterology, 2002. **122**(1): p. 44-54.
132. **Kleessen, B., A.J. Kroesen, H.J. Buhr,** *et al.*, Mucosal and invading bacteria in patients with inflammatory bowel disease compared with controls. Scand J Gastroenterol, 2002. **37**(9): p. 1034-41.
133. **Andoh, A., S. Sakata, Y. Koizumi,** *et al.*, Terminal restriction fragment length polymorphism analysis of the diversity of fecal microbiota in patients with ulcerative colitis. Inflamm Bowel Dis, 2007. **13**(8): p. 955-62.
134. **Noor, S.O., K. Ridgway, L. Scovell,** *et al.*, Ulcerative colitis and irritable bowel patients exhibit distinct abnormalities of the gut microbiota. BMC Gastroenterol, 2010. **10**: p. 134.
135. **Takaishi, H., T. Matsuki, A. Nakazawa,** *et al.*, Imbalance in intestinal microflora constitution could be involved in the pathogenesis of inflammatory bowel disease. International Journal of Medical Microbiology, 2008. **298**(5–6): p. 463-472.
136. **Lepage, P., M.C. Leclerc, M. Joossens,** *et al.*, A metagenomic insight into our gut's microbiome. Gut, 2012. **62**(1):146-58
137. **Nagalingam, N.A. and S.V. Lynch,** Role of the microbiota in inflammatory bowel diseases. Inflamm Bowel Dis, 2012. **18**(5): p. 968-84.
138. **Martinez, C., M. Antolin, J. Santos,** *et al.*, Unstable composition of the fecal microbiota in ulcerative colitis during clinical remission. Am J Gastroenterol, 2008. **103**(3): p. 643-8.
139. **Riera-Oliver, J.**, *Definiciones conceptuales de la enfermedad inflamatoria intestinal. Concepto y cronicidad.* Second ed. Enfermedad Inflamatoria Intestinal, ed. G.F. Edited by Gasull MA, Hinojosa J, Obrador A, . 2002, Madrid: Ediciones Ergón. 3-5.
140. **Lepage, P., R. Hasler, M.E. Spehlmann,** *et al.*, Twin study indicates loss of interaction between microbiota and mucosa of patients with ulcerative colitis. Gastroenterology, 2011. **141**(1): p. 227-36.
141. **Pitcher, M.C., E.R. Beatty, and J.H. Cummings,** The contribution of sulphate reducing bacteria and 5-aminosalicylic acid to faecal sulphide in patients with ulcerative colitis. Gut, 2000. **46**(1): p. 64-72.
142. **Linskens, R.K., X.W. Huijsdens, P.H. Savelkoul,** *et al.*, The bacterial flora in inflammatory bowel disease: current insights in pathogenesis and the influence of antibiotics and probiotics. Scand J Gastroenterol Suppl, 2001 (234): p. 29-40.
143. **Wang, M., G. Molin, S. Ahrne,** *et al.,* High proportions of proinflammatory bacteria on the colonic mucosa in a young patient with ulcerative colitis as revealed by cloning and sequencing of 16S rRNA genes. Dig Dis Sci, 2007. **52**(3): p. 620-7.
144. **Louis, E., A. Collard, A.F. Oger,** *et al.*, Behaviour of Crohn's disease according to the Vienna classification: changing pattern over the course of the disease. Gut, 2001. **49**(6): p. 777-82.
145. **Stange, E.F., S.P. Travis, S. Vermeire,** *et al.*, European evidence based consensus on the diagnosis and management of Crohn's disease: definitions and diagnosis. Gut, 2006. **55** Suppl 1: p. i1-15.
146. **Geboes, K.**, *Anatomía patológica de la enfermedad inflamatoria intestinal.* Third ed. Enfermedad Inflamatoria Intestinal, ed. G.F. Edited by Gasull MA, Hinojosa J, Obrador A. 2007, Madrid: Arán Ediciones S.L.
147. **Gisbert, J.**, *Marcadores biológicos y pronósticos*. Enfermedad Inflamatoria Intestinal, ed. G.F. Edited by Gasull MA, Hinojosa J, Obrador A. 2007, Madrid: Arán Ediciones S.L.
148. **Benkhadra, F. and R.L. Humbel,** Les marqueurs sérologiques des maladies inflammatoires chroniques intestinales (MICI). Immuno-analyse & Biologie Spécialisée, 2008. **23**(4): p. 202-211.
149. **Brugere, J.F., A. Mihajlovski, M. Missaoui,** *et al.*, Tools for stools: the challenge of assessing human intestinal microbiota using molecular diagnostics. Expert Rev Mol Diagn, 2009. **9**(4): p. 353-65.
150. **Andoh, A., H. Kuzuoka, T. Tsujikawa,** *et al.*, Multicenter analysis of fecal microbiota profiles in Japanese patients with Crohn's disease. J Gastroenterol, 2012. **47**(12): p. 1298-307.
151. **Williams, H.R., I.J. Cox, D.G. Walker,** *et al.*, Characterization of inflammatory bowel disease with urinary metabolic profiling. Am J Gastroenterol, 2009. **104**(6): p. 1435-44.
152. **Baumgart, D.C. and W.J. Sandborn**, Inflammatory bowel disease: clinical aspects and established and evolving therapies. Lancet, 2007. **369**(9573): p. 1641-57.
153. **Garcia-Bosch, O., E. Ricart, and J. Panes**, Review article: stem cell therapies for inflammatory bowel disease - efficacy and safety. Aliment Pharmacol Ther, 2010. **32**(8): p. 939-52.
154. **Panes, J., I. Ordas, and E. Ricart**, Stem cell treatment for Crohn's disease. Expert Rev Clin Immunol, 2010. **6**(4): p. 597-605.
155. **Salas, A., E. Ricart, and J. Panes**, Cell therapies for inflammatory bowel diseases. Expert Rev Gastroenterol Hepatol, 2009. **3**(4): p. 321-4.
156. **Srikanth, C.V. and B.A. McCormick**, Interactions of the intestinal epithelium with the pathogen and the indigenous microbiota: a three-way crosstalk. Interdiscip Perspect Infect Dis, 2008. 2008: p. 626827.

157. **Shanahan, F.,** Probiotics in inflamatory bowel disease. Gut, 2001. **48**(5): p. 609.
158. **Seksik, P., X. Dray, H. Sokol,** *et al.*, Is there any place for alimentary probiotics, prebiotics or synbiotics, for patients with inflammatory bowel disease? Mol Nutr Food Res, 2008. **52**(8): p. 906-12.
159. **Pouillart, P.R., F. Depeint, A. Abdelnour,** *et al.*, Nutriose, a prebiotic low-digestible carbohydrate, stimulates gut mucosal immunity and prevents TNBS-induced colitis in piglets. Inflamm Bowel Dis, 2010. **16**(5): p. 783-94.
160. **Mazmanian, S.K., J.L. Round, and D.L. Kasper,** A microbial symbiosis factor prevents intestinal inflammatory disease. Nature, 2008. **453**(7195): p. 620-5.
161. **Makharia, G.K., A.K. Verma, R. Amarchand,** *et al.*, Prevalence of irritable bowel syndrome: a community based study from northern India. J Neurogastroenterol Motil, 2011. **17**(1): p. 82-7.
162. **Lacy, B.E. and R.D. Lee**, Irritable bowel syndrome: a syndrome in evolution. J Clin Gastroenterol, 2005. **39**(5 Suppl 3): p. S230-42.
163. **Posserud, I., A. Ersryd, and M. Simren**, Functional findings in irritable bowel syndrome. World J Gastroenterol, 2006. **12**(18): p. 2830-8.
164. **Thompson, W.G.,** The road to rome. Gastroenterology, 2006. **130**(5): p. 1552-6.
165. **Agrawal, A. and P.J. Whorwell,** Irritable bowel syndrome: diagnosis and management. BMJ, 2006. **332**(7536): p. 280-3.
166. **Ghoshal, U.C., R. Shukla, U. Ghoshal,** *et al.*, The gut microbiota and irritable bowel syndrome: friend or foe? Int J Inflam, 2012. **2012**: p. 151085.
167. **Drossman, D.A.**, What does the future hold for irritable bowel syndrome and the functional gastrointestinal disorders? J Clin Gastroenterol, 2005. **39**(5 Suppl 3): p. S251-6.
168. **Olbe, L.**, Concept of Crohn's disease being conditioned by four main components, and irritable bowel syndrome being an incomplete Crohn's disease. Scand J Gastroenterol, 2008. **43**(2): p. 234-41.
169. **Kassinen, A., L. Krogius-Kurikka, H. Makivuokko,** *et al.*, The fecal microbiota of irritable bowel syndrome patients differs significantly from that of healthy subjects. Gastroenterology, 2007. **133**(1): p. 24-33.
170. **Whitehead, W.E., B.T. Engel, and M.M. Schuster**, Irritable bowel syndrome: physiological and psychological differences between diarrhea-predominant and constipation-predominant patients. Dig Dis Sci., 1980. **25**(6): p. 404-13.
171. **Kerckhoffs, A.P., K. Ben-Amor, M. Samsom,** *et al.*, Molecular analysis of faecal and duodenal samples reveals significantly higher prevalence and numbers of Pseudomonas aeruginosa in irritable bowel syndrome. J Med Microbiol, 2011. **60**(Pt 2): p. 236-45.
172. **Simren, M. and P.O. Stotzer,** Use and abuse of hydrogen breath tests. Gut, 2006. **55**(3): p. 297-303.
173. **Matto, J., L. Maunuksela, K. Kajander,** *et al.*, Composition and temporal stability of gastrointestinal microbiota in irritable bowel syndrome--a longitudinal study in IBS and control subjects. FEMS Immunol Med Microbiol, 2005. **43**(2): p. 213-22.
174. **Maukonen, J., R. Satokari, J. Matto,** *et al.*, Prevalence and temporal stability of selected clostridial groups in irritable bowel syndrome in relation to predominant faecal bacteria. J Med Microbiol, 2006. **55**(Pt 5): p. 625-33.
175. **Ponnusamy, K., J.N. Choi, J. Kim,** *et al.*, Microbial community and metabolomic comparison of irritable bowel syndrome faeces. J Med Microbiol, 2011. **60**(Pt 6): p. 817-27.
176. **Rajilic-Stojanovic, M., E. Biagi, H.G. Heilig,** *et al.*, Global and deep molecular analysis of microbiota signatures in fecal samples from patients with irritable bowel syndrome. Gastroenterology, 2011. **141**(5): p. 1792-801.
177. **Chadwick, V.S., W. Chen, D. Shu,** *et al.*, Activation of the mucosal immune system in irritable bowel syndrome. Gastroenterology, 2002. **122**(7): p. 1778-1783.
178. **Louis, P., G.L. Hold, and H.J. Flint**, The gut microbiota, bacterial metabolites and colorectal cancer. Nat Rev Microbiol, 2014. **12**(10): p. 661-72.
179. **Lupp, C., M.L. Robertson, M.E. Wickham,** *et al.*, Host-mediated inflammation disrupts the intestinal microbiota and promotes the overgrowth of Enterobacteriaceae. Cell Host Microbe, 2007. **2**(3): p. 204.
180. **Stecher, B. and W.D. Hardt**, The role of microbiota in infectious disease. Trends Microbiol, 2008. **16**(3): p. 107-14.
181. **Poschl, G. and H.K. Seitz**, Alcohol and cancer. Alcohol Alcohol, 2004. **39**(3): p. 155-65.
182. **Wang, X. and M.M. Huycke,** Extracellular superoxide production by *Enterococcus faecalis* promotes chromosomal instability in mammalian cells. Gastroenterology, 2007. **132**(2): p. 551-61.
183. **Bingham, S.A., B. Pignatelli, J.R. Pollock,** *et al.*, Does increased endogenous formation of N-nitroso compounds in the human colon explain the association between red meat and colon cancer? Carcinogenesis, 1996. **17**(3): p. 515-23.
184. **Medani, M., D. Collins, N.G. Docherty,** *et al.*, Emerging role of hydrogen sulfide in colonic physiology and pathophysiology. Inflamm Bowel Dis, 2011. **17**(7): p. 1620-5.

185. **Huycke, M.M., V. Abrams, and D.R. Moore**, *Enterococcus faecalis* produces extracellular superoxide and hydrogen peroxide that damages colonic epithelial cell DNA. Carcinogenesis, 2002. **23**(3): p. 529-36.
186. **Gill, C.I. and I.R. Rowland**, Diet and cancer: assessing the risk. Br J Nutr, 2002. **88** Suppl 1: p. S73-87.
187. **Balamurugan, R., E. Rajendiran, S. George,** *et al.*, Real-time polymerase chain reaction quantification of specific butyrate-producing bacteria, *Desulfovibrio* and *Enterococcus faecalis* in the feces of patients with colorectal cancer. J Gastroenterol Hepatol, 2008. **23**(8 Pt 1): p. 1298-303.
188. **Ou, J., F. Carbonero, E.G. Zoetendal,** *et al.*, Diet, microbiota, and microbial metabolites in colon cancer risk in rural Africans and African Americans. Am J Clin Nutr, 2013. **98**(1): p. 111-20.
189. **Wu, N., X. Yang, R. Zhang,** *et al.*, Dysbiosis signature of fecal microbiota in colorectal cancer patients. Microb Ecol, 2013. **66**(2): p. 462-70.
190. **Sobhani, I., J. Tap, F. Roudot-Thoraval,** *et al.*, Microbial dysbiosis in colorectal cancer (CRC) patients. PLoS One, 2011. **6**(1): p. e16393.
191. **Moore, W.E. and L.H. Moore,** Intestinal floras of populations that have a high risk of colon cancer. Appl. Environ. Microbiol., 1995. **61**(9): p. 3202-7.
192. **Ellmerich, S., M. Scholler, B. Duranton,** *et al.*, Promotion of intestinal carcinogenesis by Streptococcus bovis. Carcinogenesis, 2000. **21**(4): p. 753-6.
193. **Buc, E., D. Dubois, P. Sauvanet**, *et al.*, High prevalence of mucosa-associated E. coli producing cyclomodulin and genotoxin in colon cancer. PLoS One, 2013. **8**(2): p. e56964.
194. **Moore, W.E.C. and L.V. Holdeman,** New Names and Combinations in the Genera *Bacteroides* Castellani and Chalmers, *Fusobacterium* Knorr, *Eubacterium* Prévot, *Propionibacterium* Delwich, and *Lactobacillus* Orla-Jensen. International Journal of Systematic Bacteriology, 1973. **23**(1): p. 69-74.
195. **Cato, E.P., C.W. Salmon, and W.E.C. Moore**, *Fusobacterium prausnitzii* (Hauduroy *et al.*) Moore and Holdeman: Emended Description and Designation of Neotype Strain. International Journal of Systematic Bacteriology, 1974. **24**(2): p. 225-229.
196. **Wang, R.F., W.W. Cao, and C.E. Cerniglia,** Phylogenetic analysis of *Fusobacterium prausnitzii* based upon the 16S rRNA gene sequence and PCR confirmation. Int J Syst Bacteriol, 1996. **46**(1): p. 341-3.
197. **Duncan, S.H., G.L. Hold, H.J. Harmsen,** *et al.*, Growth requirements and fermentation products of *Fusobacterium prausnitzii*, and a proposal to reclassify it as *Faecalibacterium prausnitzii* gen. nov., comb. nov. Int. J. Syst. Evol. Microbiol. , 2002. **52**(Pt 6): p. 2141-6.
198. **Suau, A., V. Rochet, A. Sghir,** *et al.*, *Fusobacterium prausnitzii* and related species represent a dominant group within the human fecal flora. Systematic and Applied Microbiology, 2001. **24**(1): p. 139-145.
199. **Lay, C., M. Sutren, V. Rochet,** *et al.*, Design and validation of 16S rRNA probes to enumerate members of the Clostridium leptum subgroup in human faecal microbiota. Environ Microbiol, 2005. **7**(7): p. 933-46.
200. **Miyazaki, K., J.C. Martin, R. Marinsek-Logar,** *et al.*, Degradation and utilization of xylans by the rumen anaerobe *Prevotella bryantii* (formerly *P. ruminicola* subsp. *brevis*) B(1)4. Anaerobe, 1997. **3**(6): p. 373-81.
201. **Prausnitz, C.**, Der Bacillus mucosus anaerobius. Zentralbl. Bakteriol. Parasitenkd. Infektionskr. Hyg Abt. I Orig. , 1922. **89**: p. 126-132.
202. **Barcenilla, A., S.E. Pryde, J.C. Martin,** *et al.*, Phylogenetic relationships of butyrate-producing bacteria from the human gut. Appl. Environ. Microbiol., 2000. **66**(4): p. 1654-61.
203. **Louis, P., S.H. Duncan, S.I. McCrae,** *et al.*, Restricted distribution of the butyrate kinase pathway among butyrate-producing bacteria from the human colon. J. Bacteriol., 2004. **186**(7): p. 2099-106.
204. **Khan, M.T., S.H. Duncan, A.J. Stams,** *et al.*, The gut anaerobe *Faecalibacterium prausnitzii* uses an extracellular electron shuttle to grow at oxic-anoxic interphases. ISME J, 2012. **6**(8): p. 1578-85.
205. **Khan, M.T., J.M. van Dijl, and H.J. Harmsen**, Antioxidants keep the potentially probiotic but highly oxygen-sensitive human gut bacterium *Faecalibacterium prausnitzii* alive at ambient air. PLoS One, 2014. **9**(5): p. e96097.
206. **Jia, W., R.N. Whitehead, L. Griffiths,** *et al.*, Is the abundance of *Faecalibacterium prausnitzii* relevant to Crohn's disease? FEMS Microbiol. Lett., 2010. **310**(2): p. 138-44.
207. **Hold, G.L., A. Schwiertz, R.I. Aminov,** *et al.*, Oligonucleotide probes that detect quantitatively significant groups of butyrate-producing bacteria in human feces. Appl Environ Microbiol, 2003. **69**(7): p. 4320-4.
208. **Rinttila, T., A. Kassinen, E. Malinen,** *et al.*, Development of an extensive set of 16S rDNA-targeted primers for quantification of pathogenic and indigenous bacteria in faecal samples by real-time PCR. J Appl Microbiol, 2004. **97**(6): p. 1166-77.
209. **Bartosch, S., A. Fite, G.T. Macfarlane**, *et al.*, Characterization of bacterial communities in feces from healthy elderly volunteers and hospitalized elderly patients by using real-time PCR and effects of antibiotic treatment on the fecal microbiota. Appl Environ Microbiol, 2004. **70**(6): p. 3575-81.

210. **Feng, J., W. Hua, L. Zhao,** *et al.*, Specificity comparison of three *Feacalibacterium prausnitzii*-specific PCR primer pairs. Wei Sheng Wu Xue Bao, 2011. **51**(6): p. 819-27.
211. **Nava, G.M. and T.S. Stappenbeck,** Diversity of the autochthonous colonic microbiota. Gut Microbes, 2011. **2**(2): p. 99-104.
212. **Macfarlane, S. and G.T. Macfarlane,** Composition and metabolic activities of bacterial biofilms colonizing food residues in the human gut. Appl Environ Microbiol, 2006. **72**(9): p. 6204-11.
213. **Archer, S., S. Meng, J. Wu,** *et al.*, Butyrate inhibits colon carcinoma cell growth through two distinct pathways. Surgery, 1998. **124**(2): p. 248-53.
214. **Christl, S.U., H.-D. Eisner, G. Dusel,** *et al.*, Antagonistic effects of sulfide and butyrate on proliferation of colonic mucosa. Digestive Diseases and Sciences, 1996. **41**(12): p. 2477-2481.
215. **Inan, M.S., R.J. Rasoulpour, L. Yin,** *et al.*, The luminal short-chain fatty acid butyrate modulates NF-kappaB activity in a human colonic epithelial cell line. Gastroenterology, 2000. **118**(4): p. 724-34.
216. **Schwab, M., V. Reynders, S. Loitsch,** *et al.*, Involvement of different nuclear hormone receptors in butyrate-mediated inhibition of inducible NF kappa B signalling. Mol Immunol, 2007. **44**(15): p. 3625-32.
217. **Klampfer, L., J. Huang, T. Sasazuki,** *et al.*, Inhibition of interferon gamma signaling by the short chain fatty acid butyrate. Mol Cancer Res, 2003. **1**(11): p. 855-62.
218. **Qiu, X., M. Zhang, X. Yang,** *et al.*, *Faecalibacterium prausnitzii* upregulates regulatory T cells and anti-inflammatory cytokines in treating TNBS-induced colitis. J Crohns Colitis, 2013. **7**(11): p. e558-68.
219. **Martin, R., F. Chain, S. Miquel,** *et al.*, The Commensal Bacterium *Faecalibacterium prausnitzii* Is Protective in DNBS-induced Chronic Moderate and Severe Colitis Models. Inflamm Bowel Dis, 2014. **20**(3):417-30.
220. **Martin, R., S. Miquel, F. Chain,** *et al.*, *Faecalibacterium prausnitzii* prevents physiological damages in a chronic low-grade inflammation murine model. BMC Microbiol, 2015. **15**: p. 67.
221. **Quevrain, E., M.A. Maubert, C. Michon,** *et al.*, Identification of an anti-inflammatory protein from *Faecalibacterium prausnitzii*, a commensal bacterium deficient in Crohn's disease. Gut, 2015. doi: 10.1136/gutjnl-2014-307649.
222. **Carlsson, A.H., O. Yakymenko, I. Olivier,** *et al.*, *Faecalibacterium prausnitzii* supernatant improves intestinal barrier function in mice DSS colitis. Scand J Gastroenterol, 2013. **48**(10): p. 1136-44.
223. **Wrzosek, L., S. Miquel, M.L. Noordine,** *et al.*, *Bacteroides thetaiotaomicron* and *Faecalibacterium prausnitzii* influence the production of mucus glycans and the development of goblet cells in the colonic epithelium of a gnotobiotic model rodent. BMC Biol, 2013. **11**: p. 61.
224. **Ohman, L. and M. Simren,** New insights into the pathogenesis and pathophysiology of irritable bowel syndrome. Dig Liver Dis, 2007. **39**(3): p. 201-15.
225. **Miquel, S., R. Martin, O. Rossi,** *et al.*, *Faecalibacterium prausnitzii* and human intestinal health. Curr Opin Microbiol, 2013. 16(3): p. 255-61.
226. **Li, M., B. Wang, M. Zhang,** *et al.*, Symbiotic gut microbes modulate human metabolic phenotypes. Proc Natl Acad Sci U S A, 2008. **105**(6): p. 2117-22.
227. **Jansson, J., B. Willing, M. Lucio,** *et al.*, Metabolomics reveals metabolic biomarkers of Crohn's disease. PLoS One, 2009. **4**(7): p. e6386.
228. **Miquel, S., M. Leclerc, R. Martin,** *et al.*, Identification of metabolic signatures linked to anti-inflammatory effects of Faecalibacterium prausnitzii. MBio, 2015. **6**(2).
229. **Castillo, M., G. Skene, M. Roca, et al**., Application of 16S rRNA gene-targetted fluorescence in situ hybridization and restriction fragment length polymorphism to study porcine microbiota along the gastrointestinal tract in response to different sources of dietary fibre. FEMS Microbiol Ecol, 2007. **59**(1): p. 138-46.
230. **Nava, G.M., H.J. Friedrichsen, and T.S. Stappenbeck**, Spatial organization of intestinal microbiota in the mouse ascending colon. ISME J, 2011. **5**(4): p. 627-38.
231. **Oikonomou, G., A.G. Teixeira, C. Foditsch,** *et al.*, Fecal microbial diversity in pre-weaned dairy calves as described by pyrosequencing of metagenomic 16S rDNA. Associations of *Faecalibacterium* species with health and growth. PLoS One, 2013. **8**(4): p. e63157.
232. **Bjerrum, L., R.M. Engberg, T.D. Leser,** *et al.*, Microbial community composition of the ileum and cecum of broiler chickens as revealed by molecular and culture-based techniques. Poult Sci, 2006. **85**(7): p. 1151-64.
233. **Scupham, A.J**., Succession in the intestinal microbiota of preadolescent turkeys. FEMS Microbiol Ecol, 2007. **60**(1): p. 136-47.
234. **Hopkins, M.J., G.T. Macfarlane, E. Furrie,** *et al.*, Characterisation of intestinal bacteria in infant stools using real-time PCR and northern hybridisation analyses. FEMS Microbiol Ecol, 2005. **54**(1): p. 77-85.
235. **Lozupone, C., K. Faust, J. Raes,** *et al.*, Identifying genomic and metabolic features that can underlie early successional and opportunistic lifestyles of human gut symbionts. Genome Res, 2012. **22**(10): p. 1974-84.

236. **Matsuki, T., K. Watanabe, J. Fujimoto,** *et al.*, Use of 16S rRNA gene-targeted group-specific primers for real-time PCR analysis of predominant bacteria in human feces. Appl Environ Microbiol, 2004. **70**(12): p. 7220-8.
237. **Schwiertz, A., M. Jacobi, J.S. Frick,** *et al.*, Microbiota in pediatric inflammatory bowel disease. J Pediatr, 2010. **157**(2): p. 240-244 e1.
238. **van Tongeren, S.P., J.P. Slaets, H.J. Harmsen,** *et al.*, Fecal microbiota composition and frailty. Appl. Environ. Microbiol., 2005. **71**(10): p. 6438-6442.
239. **Mariat, D., O. Firmesse, F. Levenez,** *et al.*, The *Firmicutes/Bacteroidetes* ratio of the human microbiota changes with age. BMC Microbiol., 2009. **9**: p. 123-128.
240. **Salazar, N., M. Gueimonde, A.M. Hernandez-Barranco,** *et al.*, Exopolysaccharides produced by intestinal Bifidobacterium strains act as fermentable substrates for human intestinal bacteria. Appl Environ Microbiol, 2008. **74**(15): p. 4737-45.
241. **Hansen, R., R.K. Russell, C. Reiff,** *et al.*, Microbiota of de-novo pediatric IBD: increased Faecalibacterium prausnitzii and reduced bacterial diversity in Crohn's but not in ulcerative colitis. Am J Gastroenterol, 2012. **107**(12): p. 1913-22.
242. **Furet, J.P., L.C. Kong, J. Tap,** *et al.*, Differential adaptation of human gut microbiota to bariatric surgery-induced weight loss: links with metabolic and low-grade inflammation markers. Diabetes, 2010. **59**(12): p. 3049-57.
243. **McLaughlin, S.D., S.K. Clark, P.P. Tekkis,** *et al.*, The bacterial pathogenesis and treatment of pouchitis. Therap Adv Gastroenterol, 2010. **3**(6): p. 335-48.
244. **Vermeiren, J., P. Van den Abbeele, D. Laukens,** *et al.*, Decreased colonization of fecal Clostridium coccoides/Eubacterium rectale species from ulcerative colitis patients in an in vitro dynamic gut model with mucin environment. FEMS Microbiol Ecol, 2012. **79**(3): p. 685-96.
245. **De Palma, G., I. Nadal, M.C. Collado,** *et al.*, Effects of a gluten-free diet on gut microbiota and immune function in healthy adult human subjects. British Journal of Nutrition, 2009. **102**(08): p. 1154-1160.
246. **Fujimoto, T., H. Imaeda, K. Takahashi,** *et al.*, Decreased abundance of Faecalibacterium prausnitzii in the gut microbiota of Crohn's disease. J Gastroenterol Hepatol, 2013. **28**(4): p. 613-9.
247. **Dorffel, Y., A. Swidsinski, V. Loening-Baucke,** *et al.*, Common biostructure of the colonic microbiota in neuroendocrine tumors and Crohn's disease and the effect of therapy. Inflamm Bowel Dis, 2012. **18**(9): p. 1663-71.
248. **Ng, S.C., J.L. Benjamin, N.E. McCarthy,** *et al.*, Relationship between human intestinal dendritic cells, gut microbiota, and disease activity in Crohn's disease. Inflamm Bowel Dis, 2011. **17**(10): p. 2027-37.
249. **Benjamin, J.L., C.R. Hedin, A. Koutsoumpas,** *et al.*, Smokers with active Crohn's disease have a clinically relevant dysbiosis of the gastrointestinal microbiota. Inflamm Bowel Dis, 2012. **18**(6): p. 1092-100.
250. **Wang, W., L. Chen, R. Zhou,** *et al.*, Increased proportions of *Bifidobacterium* and the *Lactobacillus* group and loss of butyrate-producing bacteria in inflammatory bowel disease. J Clin Microbiol, 2014. **52**(2): p. 398-406.
251. **Li, Q., C. Wang, C. Tang,** *et al.*, Molecular-phylogenetic characterization of the microbiota in ulcerated and non-ulcerated regions in the patients with Crohn's disease. PLoS One, 2012. **7**(4): p. e34939.
252. **Varela, E., C. Manichanh, M. Gallart,** *et al.*, Colonisation by *Faecalibacterium prausnitzii* and maintenance of clinical remission in patients with ulcerative colitis. Aliment Pharmacol Ther, 2013. **38**(2): p. 151-61.
253. **Rigsbee, L., R. Agans, V. Shankar,** *et al.*, Quantitative profiling of gut microbiota of children with diarrhea-predominant irritable bowel syndrome. Am J Gastroenterol, 2012. **107**(11): p. 1740-51.
254. **Duboc, H., D. Rainteau, S. Rajca,** *et al.*, Increase in fecal primary bile acids and dysbiosis in patients with diarrhea-predominant irritable bowel syndrome. Neurogastroenterol Motil, 2012. **24**(6): p. 513-20.
255. **Balamurugan, R., G. George, J. Kabeerdoss,** *et al.*, Quantitative differences in intestinal Faecalibacterium prausnitzii in obese Indian children. Br J Nutr, 2010. **103**(3): p. 335-8.
256. **Feng, J., H. Tang, M. Li,** *et al.*, The abundance of fecal *Faecalibacterium prausnitzii* in relation to obesity and gender in Chinese adults. Arch Microbiol, 2014. **196**(1): p. 73-7.
257. **Payne, A.N., C. Chassard, M. Zimmermann,** *et al.*, The metabolic activity of gut microbiota in obese children is increased compared with normal-weight children and exhibits more exhaustive substrate utilization. Nutr Diabetes, 2011. **1**: p. e12.
258. **Graessler, J., Y. Qin, H. Zhong,** *et al.*, Metagenomic sequencing of the human gut microbiome before and after bariatric surgery in obese patients with type 2 diabetes: correlation with inflammatory and metabolic parameters. Pharmacogenomics J, 2012.

259. **Swidsinski, A., Y. Dorffel, V. Loening-Baucke,** *et al.*, Acute appendicitis is characterised by local invasion with Fusobacterium nucleatum/necrophorum. Gut, 2011. **60**(1): p. 34-40.
260. **Gerasimidis, K., M. Bertz, L. Hanske,** *et al.*, Decline in presumptively protective gut bacterial species and metabolites are paradoxically associated with disease improvement in pediatric Crohn's disease during enteral nutrition. Inflamm Bowel Dis, 2014. **20**(5): p. 861-71.
261. **Dewulf, E.M., P.D. Cani, S.P. Claus,** *et al.*, Insight into the prebiotic concept: lessons from an exploratory, double blind intervention study with inulin-type fructans in obese women. Gut, 2013. **62**(8): p. 1112-21.
262. **Benus, R.F., H.J. Harmsen, G.W. Welling,** *et al.*, Impact of digestive and oropharyngeal decontamination on the intestinal microbiota in ICU patients. Intensive Care Med, 2010. **36**(8): p. 1394-402.
263. **Maccaferri, S., B. Vitali, A. Klinder,** *et al.*, Rifaximin modulates the colonic microbiota of patients with Crohn's disease: an *in vitro* approach using a continuous culture colonic model system. J Antimicrob Chemother, 2010. **65**(12): p. 2556-65.
264. **Foditsch, C., T.M. Santos, A.G. Teixeira,** *et al.*, Isolation and characterization of *Faecalibacterium prausnitzii* from calves and piglets. PLoS One, 2014. **9**(12): p. e116465.
265. **Pryde, S.E., S.H. Duncan, G.L. Hold,** *et al.*, The microbiology of butyrate formation in the human colon. FEMS Microbiol Lett, 2002. 217(2): p. 133-9.
266. **Leser, T.D., J.Z. Amenuvor, T.K. Jensen,** *et al.*, Culture-independent analysis of gut bacteria: the pig gastrointestinal tract microbiota revisited. Appl Environ Microbiol, 2002. **68**(2): p. 673-90.
267. **Pryde, S.E., A.J. Richardson, C.S. Stewart,** *et al.*, Molecular analysis of the microbial diversity present in the colonic wall, colonic lumen, and cecal lumen of a pig. Appl Environ Microbiol, 1999. **65**(12): p. 5372-7.
268. **de Goffau, M.C., K. Luopajarvi, M. Knip,** *et al.*, Fecal microbiota composition differs between children with beta-cell autoimmunity and those without. Diabetes, 2013. **62**(4): p. 1238-44.
269. **Karlsson, F.H., V. Tremaroli, I. Nookaew,** *et al.*, Gut metagenome in European women with normal, impaired and diabetic glucose control. Nature, 2013. **498**(7452): p. 99-103.
270. **Louis, P., K.P. Scott, S.H. Duncan,** *et al.*, Understanding the effects of diet on bacterial metabolism in the large intestine. J Appl Microbiol, 2007. **102**(5): p. 1197-208.
271. **Licht, T., M. Hansen, A. Bergstrom,** *et al.*, Effects of apples and specific apple components on the cecal environment of conventional rats: role of apple pectin. BMC Microbiol., 2010. **10**(1): p. 13-23.
272. **Onumpai, C., S. Kolida, E. Bonnin,** *et al.*, Microbial utilization and selectivity of pectin fractions with various structures. Appl Environ Microbiol, 2011. **77**(16): p. 5747-54.
273. **Dongowski, G., A. Lorenz, and H. Anger,** Degradation of pectins with different degrees of esterification by Bacteroides thetaiotaomicron isolated from human gut flora. Appl Environ Microbiol, 2000. **66**(4): p. 1321-7.
274. **Salvatore, S., R. Heuschkel, S. Tomlin,** *et al.*, A pilot study of N-acetyl glucosamine, a nutritional substrate for glycosaminoglycan synthesis, in paediatric chronic inflammatory bowel disease. Aliment Pharmacol Ther, 2000. **14**(12): p. 1567-79.
275. **Duncan, S.H., G. Holtrop, G.E. Lobley,** *et al.*, Contribution of acetate to butyrate formation by human faecal bacteria. Br. J. Nutr., 2004. **91**(06): p. 915-923.
276. **Willis, C.L., J.H. Cummings, G. Neale,** *et al.*, *In vitro* effects of mucin fermentation on the growth of human colonic sulphate-reducing bacteria. Anaerobe, 1996(2): p. 117-122.
277. **Collado, M.C., M. Derrien, E. Isolauri,** *et al.*, Intestinal integrity and Akkermansia muciniphila, a mucin-degrading member of the intestinal microbiota present in infants, adults, and the elderly. Appl Environ Microbiol, 2007. **73**(23): p. 7767-70.
278. **Derrien, M., M.C. Collado, K. Ben-Amor,** *et al.*, The Mucin degrader Akkermansia muciniphila is an abundant resident of the human intestinal tract. Appl Environ Microbiol, 2008. **74**(5): p. 1646-8.
279. **Derrien, M., E.E. Vaughan, C.M. Plugge,** *et al.*, *Akkermansia muciniphila* gen. nov., sp. nov., a human intestinal mucin-degrading bacterium. Int J Syst Evol Microbiol, 2004. **54**(Pt 5): p. 1469-76.
280. **van Passel, M.W., R. Kant, E.G. Zoetendal,** *et al.*, The genome of *Akkermansia muciniphila*, a dedicated intestinal mucin degrader, and its use in exploring intestinal metagenomes. PLoS One, 2011. **6**(3): p. e16876.
281. **Looft, T., H.K. Allen, B.L. Cantarel,** *et al.*, Bacteria, phages and pigs: the effects of in-feed antibiotics on the microbiome at different gut locations. ISME J, 2014. **8**(8): p. 1566-76.
282. **Nugent, S.G., D. Kumar, D.S. Rampton,** *et al.*, Intestinal luminal pH in inflammatory bowel disease: possible determinants and implications for therapy with aminosalicylates and other drugs. Gut, 2001. **48**(4): p. 571-7.
283. **Barkas, F., E. Liberopoulos, A. Kei,** *et al.*, Electrolyte and acid-base disorders in inflammatory bowel disease. Ann Gastroenterol, 2013. **26**(1): p. 23-28.
284. **Lapidus, A. and C. Einarsson,** Bile composition in patients with ileal resection due to Crohn's disease. Inflamm Bowel Dis, 1998. **4**(2): p. 89-94.

285. **Pereira, S.P., I.M. Bain, D. Kumar,** *et al.*, Bile composition in inflammatory bowel disease: ileal disease and colectomy, but not colitis, induce lithogenic bile. Aliment Pharmacol Ther, 2003. **17**(7): p. 923-33.
286. **Fisher, K. and C. Phillips,** The ecology, epidemiology and virulence of *Enterococcus*. Microbiology, 2009. **155**(Pt 6): p. 1749-57.
287. **Begley, M., C.G. Gahan, and C. Hill**, The interaction between bacteria and bile. FEMS Microbiol Rev, 2005. **29**(4): p. 625-51.
288. **Stellwag, E.J. and P.B. Hylemon**, Purification and characterization of bile salt hydrolase from Bacteroides fragilis subsp. fragilis. Biochim Biophys Acta, 1976. **452**(1): p. 165-76.
289. **Hoffmann, T.W., H.-P. Pham, C. Bridonneau,** *et al.*, Microorganisms linked to inflammatory bowel disease-associated dysbiosis differentially impact host physiology in gnotobiotic mice. ISME J, 2015. doi: 10.1038/ismej.2015.127.
290. **Martin, H.M., B.J. Campbell, C.A. Hart,** *et al.*, Enhanced *Escherichia coli* adherence and invasion in Crohn's disease and colon cancer. Gastroenterology, 2004. **127**(1): p. 80-93.
291. **Sasaki, M., S.V. Sitaraman, B.A. Babbin,** *et al.*, Invasive *Escherichia coli* are a feature of Crohn's disease. Lab Invest, 2007. **87**(10): p. 1042-54.
292. **Rossello-Mora, R. and R. Amann,** Past and future species definitions for Bacteria and Archaea. Syst Appl Microbiol, 2015. **38**(4): p. 209-16.
293. **Ze, X., S.H. Duncan, P. Louis,** *et al.*, *Ruminococcus bromii* is a keystone species for the degradation of resistant starch in the human colon. ISME J, 2012. **6**(8): p. 1535-43.
294. **Scanlan, P.D., F. Shanahan, C. O'Mahony,** *et al.*, Culture-independent analyses of temporal variation of the dominant fecal microbiota and targeted bacterial subgroups in Crohn's disease. J Clin Microbiol, 2006. **44**(11): p. 3980-8.
295. **Wang, T., G. Cai, Y. Qiu,** *et al.*, Structural segregation of gut microbiota between colorectal cancer patients and healthy volunteers. ISME J, 2012. **6**(2): p. 320-9.
296. **Mowat, C., A. Cole, A. Windsor,** *et al.*, Guidelines for the management of inflammatory bowel disease in adults. Gut, 2011. **60**(5): p. 571-607.
297. **Papa, E., M. Docktor, C. Smillie,** *et al.*, Non-invasive mapping of the gastrointestinal microbiota identifies children with inflammatory bowel disease. PLoS One, 2012. **7**(6): p. e39242.

Glossary

16S rRNA gene (rrnA)	Small subunit ribosomal gene approximately 1500 bp long. Is a highly conserved component of the trancriptional machinery of all DNA-based life forms that consists of conserved and variable regions (V1-V9). The variable regions allow discrimination between different microorganisms [1].
Concordance	The occurrence of the same trait in both members of a pair of twins. Concordance might occur for disease as well as for behaviours such as smoking.
Diversity	A combination of richness and evenness—can be considered to be a summary statistic for community structure as membership, abundance and evenness are taken into account. [1]
Dysbiosis	Disturbances in the balance in the intestinal microbiota composition
Enterotype	Classification of an organisms based on its stable climax microbial community in the gut
Evenness	A measure of the skew in abundance of community members. Is there one dominant organism or are all evenly represented [1].
Fistulising Crohn's disease	That produces fistulae, which are an abnormal passage or communication, usually between two internal organs or leading from an internal organ to the surface of the body.
Fulminant Crohn's disease	A patient may have persistent symptoms despite appropriate treatment for moderate to severe Crohn's disease or may experience high fever, persistent vomiting, evidence of intestinal obstruction or abscess, or more severe weight loss.
Functional gastrointestinal disorder	Disorders of the digestive system in which symptoms cannot be explained by the presence of structural or tissue abnormality, based on clinical symptoms.
Gnotobiotic model	A gnotobiotic animal model is an animal in which only certain known strains of bacteria and other microorganisms are present.
Ileal pouch	An artificial rectum surgically created out of ileal gut tissue in patients who have undergone a colectomy, which is created in the management of patients with ulcerative colitis and indeterminate colitis
Ileostomy	Surgical opening constructed by bringing the end or loop of small intestine (the ileum) out onto the surface of the skin.
Incidence	In epidemiology, the incidence a measure of the number of new cases that acquire a disease per unit of person and time
Metagenomics	Describes the direct sequencing of the total DNA extracted from a microbial community [1].
Microbiome	Literally means "small biome", the ecosystem comprising all microorganisms in a particular environment together with their genes and environmental interactions [1]. Genome of human microbial symbionts.
OTU	Stands for Operational Taxonomic Unit. This term arise with the use of molecular methods to study bacterial diversity, and refers to clusters of similar sequences since molecular methods for determining the identity of Bacteria and Archaea do not map directly to the classic biochemically determined taxonomies [1].
Pathognomonic	Characteristic for a particular disease (i.e. a pathognomonic sign is a particular sign whose presence means that a particular disease is present beyond any doubt).
Pouchitis	Inflammation of the ileal pouch
Prevalence	In epidemiology, total number of cases with a certain disease in the population at a given time.
Psychosomatic	Of or relating to a disorder having physical symptoms but originating from mental or emotional causes
Richness	The number of different types of organism present [1].
Refractory	Cases or patients unresponsive to any of the conventional treatments for a given condition.
Stenosis	Narrowing or stricture of a duct or canal.
Prevalence	As used herein refers to a measure of the number of cases of disease occurring within the population under study, i.e., % of biological samples or individuals positive for a target microorganism from the total of biological samples or individuals analyzed. Prevalence is thus calculated from the qualitative determination (presence/absence) of said target microorganism within each of the samples or individuals under study.
Abundance	Refers to a measure of the quantity of a target microorganism within a biological sample. It is also referred as "load". Bacterial quantification is generally carried out by molecular methods, typically by determining the number of 16S rRNA gene copies of said target microorganism, for instance by fluorescence in situ hybridization (FISH), quantitative polymerase chain reaction (qPCR) or PCR/pyrosequencing. Quantification of the abundance of a target nucleic acid sequence within a biological sample might be absolute or relative. "Relative quantification" is generally based on one or more internal reference genes, i.e., 16S rRNA genes from reference strains, such as determination of total bacteria using universal primers and expressing the abundance of the target nucleic acid sequence as a percentage of total bacterial 16S rRNA gene copies. "Absolute quantification" gives the exact number of target molecules by comparison with DNA standards.

Sensitivity	Proportion of subjects who have the target condition (reference standard positive) and give positive test results (True positive/ (True positive + False negative)). It shows how good the test is at detecting a disease. Sensitivity ("sens") may be within the range of 0 (0%) <sens< 1 (100%) and ideally, the number of false negatives equalling zero or close to equalling zero and sensitivity equalling one (100%) or close to equalling one (100%).
Specificity	Proportion of subjects without the target condition (reference standard negative) and give negative test results (True Negative/ (True Negative + False Positive)). It shows how good the test is at identifying normal (negative) condition. Specificity ("spec") may be within the range of 0 (0%) < spec < 1 (100%) and ideally, the number of false positives equalling zero or close to equalling zero and specificity equalling one (100%) or close to equalling one (100%).
Accuracy	Proportion of true results, either true positive or true negative, in a population. It measures the degree of veracity of a screening test on a condition, i.e., how correct is the determination and exclusion of a given condition (True Negative + True Positive)/ (True Negative +True Positive +False Negative +False Positive). Accuracy ("acc") may be within the range of 0 (0%) <acc< 1 (100%) and ideally, the number of false positives equalling zero or close to equalling zero and accuracy equalling one (100%) or close to equalling one (100%).
Receiver Operating Characteristic (ROC) curves	Graphical plot that illustrates the performance of a binary classifier system as its discrimination threshold is varied. The curve is created by plotting the true positive rate against the false positive rate at various threshold settings. The true positive rate is also known as sensitivity. The false positive rate is calculated as 1 - specificity. The ROC curve is thus a way of graphically displaying the true positive rate versus the false positive rate (sensitivity vs. (1-specificity)) across a range of cut-offs and of selecting the optimal cut-off for clinical use. Accuracy expressed as the area under the ROC curve (AUC) provides a useful parameter for comparing test performance. An AUC approaching 1 indicates that the test is highly sensitive as well as highly specific whereas an AUC approaching 0.5 indicates that the test is neither sensitive nor specific. In general, a test is considered to be a suitable discriminator if the AUC is from 0.6 to 0.75, to have high discrimination capacity if the AUC is from 0.75 to 0.9 and to be an excellent discriminator if the AUC is from 0.9 to 1.
Genomovar	Which stands for distinct genomic groups that are sufficiently different to be classified as different species, but with phenotypes that do not show sufficient robust differences for discriminating them [2] Cultivated strains or groups of strains that constitute genotypic entities below or at the subspecies level, but differential phenotypic characters required for categorization as subspecies are lacking [3]
Genomospecies	Groups of cultivated organisms for which data from nucleic acid comparisons indicated that they constitute new species, but for which no distinguishing phenotypic properties have yet been found [3] (i.e. a species from the genomic point of view).[2]

References

1. Cox, M.J., W.O. Cookson, and M.F. Moffatt, Sequencing the human microbiome in health and disease. Hum Mol Genet, 2013. **22**(R1): p. R88-94.
2. Rossello-Mora, R. and R. Amann, Past and future species definitions for Bacteria and Archaea. Syst Appl Microbiol, 2015. **38**(4): p. 209-16.
3. Schloter, M., M. Lebuhn, T. Heulin, et al., Ecology and evolution of bacterial microdiversity. FEMS Microbiol Rev, 2000. **24**(5): p. 647-60.

Annex

Supplemental materials *Article I*

The following supplement accompanies the article

Cultured representatives of two major phylogroups of human colonic *Faecalibacterium prausnitzii* can utilize pectin, uronic acids and host-derived substrates for growth

Mireia Lopez-Siles, Tanweer M. Khan, Sylvia H. Duncan, Hermie J. M. Harmsen, L. Jesús Garcia-Gil, and Harry J. Flint.

Applied and Environmental Microbiology 2012; 78 (2): 420-428.

This supplement includes additional data about isolates molecular fingerprinting for selection of representative strains, and results from competition experiments for pectin fermentation at different pH values of the culture media simulating different conditions of the colon.

Supplementary figures

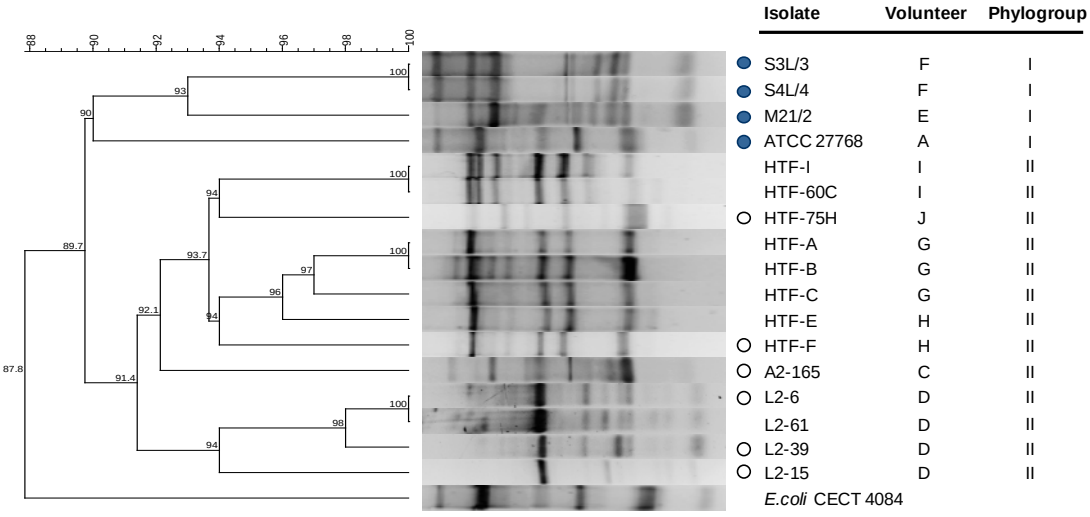


Figure S1. Hierarchical cluster of RAPD-PCR fingerprints of *F. prausnitzii* isolates by using Gelcompar II. Four phylogroup I (●) and six phylogroup II (○) isolates have been selected for further phenotypical characterisation. Similarity values are shown at branching points.
Similarity coefficient: Different bands; Dendogram type: UPGMA; Tolerance: 1%

Supplementary tables

Table S1. Short chain fatty acid (SCFA) formed by *F. prausnitzii* strains alone and in mixed culture (YcFA medium supplemented with 0.5% pectin at 3 pH values).

Initial pH 6.12						
	SCFA (mM)					
Culture(s)	Formate	Acetate	Propion.	Butyrate	Lactate	Succinate
S3L/3 only	2.4±0.2	-9.6±0.5		12.8±0.1	0.6±0.0	
A2-165 only	5.6±0.3	-6.5±1.7		9.8±0.4		
DSM3376 only	10.8±0.6	13.4±3.0			3.7±0.5	
B5482 only	0.8±0.4	3.0±1.1				8.2±0.3
S3L/3+B5482	1.9±0.2	-4.4±1.1	1.7±1.6	11.7±1.6		3.5±0.5
A2-165+B5482	3.7±0.4	3.3±2.9	1.4±0.6	6.4±0.5		5.3±0.7
3376+B5482	9.5±0.3	14.1±1.7			4.3±1.5	1.4±0.1
S3L/3+3376	8.4±0.1	2.8±1.2		10.9±0.9	2.3±0.2	
A2-165+3376	9.6±0.3	6.1±2.0		8.0±0.6	2.5±0.8	
S3L/3+B5482+3376	8.2±0.1	5.0±1.1		11.8±1.0	2.2±0.4	1.4±0.0
A2-165+B5482+3376	9.6±0.4	7.6±1.9		8.0±0.3	2.5±0.2	1.3±0.1
Initial pH 6.45						
	SCFA (mM)					
Culture(s)	Formate	Acetate	Propion.	Butyrate	Lactate	Succinate
S3L/3 only	4.4±1.0	-6.9±3.7		13.6±0.5		
A2-165 only	9.6±0.2	-0.6±0.7		10.1±0.3		
DSM3376 only	17.4±1.7	16.3±5.5			1.7±0.1	
B5482 only	0.9±0.1	16.4±1.1	3.5±0.1			11.2±0.3
S3L/3+B5482	2.6±0.4	3.2±3.4	2.1±0.4	10.4±0.4		6.1±0.4
A2-165+B5482	3.7±0.2	7.2±0.3	2.0±0.1	6.3±0.1		7.6±0.0
3376+B5482	17.6±0.4	24.1±1.2	0.2±0.1		1.1±0.1	4.3±0.2
S3L/3+3376	16.0±0.3	12.2±0.8		11.8±0.3	1.4±0.0	
A2-165+3376	17.5±1.1	13.2±3.3		7.8±0.4	1.1±0.2	
S3L/3+B5482+3376	12.8±0.1	8.9±0.4		11.9±0.3	1.1±0.0	2.9±0.1
A2-165+B5482+3376	14.8±1.4	11.7±5.2		6.9±1.0	0.9±0.3	3.1±0.1
Initial pH 6.79						
	SCFA (mM)					
Culture(s)	Formate	Acetate	Propion.	Butyrate	Lactate	Succinate
S3L/3 only	5.0±0.1	-3.7±0.5		13.2±0.5		
A2-165 only	10.3±0.7	1.1±2.8		8.7±0.3	0.4±0.4	
DSM3376 only	21.5±0.3	24.9±0.9			0.8±0.0	
B5482 only	1.0±0.4	17.9±2.6	4.2±1.2		0.7±0.1	9.9±1.1
S3L/3+B5482	3.2±0.3	14.7±1.1	2.7±0.5	9.8±0.6	0.3±0.4	7.0±0.2
A2-165+B5482	5.4±0.2	16.6±2.7	2.8±0.7	6.2±0.2		7.8±0.6
3376+B5482	15.5±0.6	24.4±1.5	2.3±0.2			7.7±0.1
S3L/3+3376	12.8±2.3	2.6±7.0		8.5±1.5	0.6±0.5	
A2-165+3376	20.0±1.5	19.3±4.2		6.9±0.5	0.7±0.1	
S3L/3+B5482+3376	13.1±2.2	9.7±3.4		8.3±2.3		2.5±2.2
A2-165+B5482+3376	11.7±2.7	9.4±2.9		6.2±1.1		3.7±0.7

Table S2. Pectin (total sugar) utilization, growth and final pH for the experiments shown in Table S1 and Figure 5 (main paper).

Initial pH 6.12				
Strain(s)	Pectin used (%)	Final OD_{650}	Final pH	Counts ($\times 10^7$/ml) *F. prausnitzii*
S3L/3 only	47.0±5.3	0.69±0.17	6.05±0.04	16.2±0.9
A2-165 only	46.8±4.2	0.70±0.02	5.80±0.02	10.6±2.1
DSM3376 only	58.9±5.9	0.76±0.02	5.20±0.02	
B5482 only	47.5±4.8	0.76±0.00	5.44±0.01	
S3L/3+B5482	57.8±4.7	1.01±0.00	5.79±0.02	24.7±0.1
A2-165+B5482	47.9±11.6	0.87±0.01	5.54±0.02	12.0±0.6
3376+B5482	60.0±2.1	0.81±0.02	5.18±0.01	
S3L/3+3376	70.0±2.2	0.98±0.01	5.44±0.02	18.9±1.3
A2-165+3376	64.8±1.2	0.92±0.02	5.37±0.02	16.4±0.3
S3L/3+B5482+3376	66.2±3.8	1.09±0.01	5.46±0.03	17.7±0.3
A2-165+B5482+3376	64.1±4.2	1.05±0.01	5.33±0.01	18.9±1.4
Initial pH 6.45				
Strain(s)	Pectin used (%)	Final OD_{650}	Final pH	Counts ($\times 10^7$/ml) *F. prausnitzii*
S3L/3 only	43.5±4.1	0.75±0.01	6.26±0.04	16.4±7.0
A2-165 only	40.4±6.1	0.76±0.01	6.15±0.03	8.8±1.4
DSM3376 only	69.4±7.7	0.93±0.02	5.61±0.02	
B5482 only	80.6±7.4	1.26±0.01	5.73±0.01	
S3L/3+B5482	70.6±3.4	1.23±0.02	5.96±0.01	24.1±1.7
A2-165+B5482	70.0±2.1	1.18±0.02	5.86±0.03	16.6±2.0
3376+B5482	60.7±4.8	1.13±0.04	5.48±0.03	
S3L/3+3376	63.0±1.4	1.14±0.01	5.85±0.03	9.4±0.6
A2-165+3376	74.7±9.5	1.07±0.01	5.73±0.02	3.9±0.0
S3L/3+B5482+3376	77.3±2.8	1.24±0.02	5.79±0.03	18.5±1.6
A2-165+B5482+3376	76.4±7.1	1.17±0.02	5.64±0.03	6.4±1.3
Initial pH 6.79				
Strain(s)	Pectin used (%)	Final OD_{650}	Final pH	Counts ($\times 10^7$/ml) *F. prausnitzii*
S3L/3 only	44.3±5.3	0.70±0.04	6.64±0.04	14.5±0.8
A2-165 only	38.3±5.9	0.75±0.03	6.67±0.08	10.1±2.2
DSM3376 only	63.7±1.9	0.89±0.01	6.22±0.03	
B5482 only	84.8±2.3	1.24±0.00	6.36±0.05	
S3L/3+B5482	83.3±6.3	1.25±0.01	6.45±0.04	8.5±1.3
A2-165+B5482	85.4±2.8	1.18±0.00	6.38±0.06	8.8±2.4
3376+B5482	83.9±3.7	1.30±0.02	5.96±0.05	
S3L/3+3376	65.8±2.9	1.09±0.02	6.26±0.02	3.3±0.1
A2-165+3376	65.7±3.4	0.99±0.02	6.18±0.09	7.7±2.5
S3L/3+B5482+3376	76.8±3.1	1.24±0.03	6.16±0.05	5.1±3.3
A2-165+B5482+3376	78.4±4.2	1.21±0.01	6.07±0.04	4.4±1.5

Supplemental materials *Article II*

The following supplement accompanies the article

Mucosa-associated *Faecalibacterium prausnitzii* population richness is reduced in inflammatory bowel disease patients

Mireia Lopez-Siles, Margarita Martinez-Medina, Carles Abellà, David Busquets, Miriam Sabat-Mir, Sylvia H. Duncan, Xavier Aldeguer, Harry J. Flint, and L. Jesús Garcia-Gil.

Applied and Environmental Microbiology, 2015; 81 (21): 7582-7592

This supplement includes additional data about primers design and OTUs equivalences at different cut-offs of similarity.

Supplementary text

Specificity tests

The specificity of the oligonucleotides was tested by comparing against the Ribosomal Database Project II (RDP) (1) and GenBank database through Seqmatch and BLAST (2) tools, respectively. The *in silico* analysis of the oligonucleotide set of choice showed that primer Fpra427F was unequivocally specific for *F. prausnitzii* and targeted all the isolates available to date whereas the Fpra1127R primer was genus-specific. The coverage of the Fpra 427F-Fpra1127R primer set was of 70.6% of the *Faecalibacterium* sequences in the SILVA dataset.

Specificity was also tested *in vitro* by testing *F. prausnitzii* DNA (10 ng) recovered from 9 isolates, representative of both phylogroups. DNAs from 71 additional representative bacterial species (see list on Table S2) which are either close relatives of *F. prausnitzii* or belong to the major groups of bacteria present in the colon were also included. The PCR reaction was carried out as described in the methods section of the main paper. Negative results were cross checked by alternative amplification with universal bacterial primers Bac27F and Uni1492R as previously reported (3, 4). Results from the specificity test are also shown in Table S2 evidencing total specificity thus successfully detecting all *F. prausnitzii* isolates. There was no cross-reaction with any of the non-target microorganisms, and negative results were validated by positive amplification by Bacteria conventional PCR with primers Bac27F and Uni1492R.

Sensitivity tests

To determine the detection limit of the assay, ten-fold serial dilutions between 10^{-4} and 10^{7} target copies of genomic DNA from *F. prausnitzii* strains ATCC27768, S3L/3, L2-6 and A2-165 were analysed as detailed in the section "PCR amplification and DGGE fingerprinting" in the main text. Data was analyzed by a Probit test (Minitab® 14 Statistical Software, Pennsylvania, USA), in which the ratio of positive/negative amplification events was plotted against the amount of target genes present per reaction. Considering all the strains, the theoretical minimum number of 16S rRNA genes of *F. prausnitzii* per reaction to have a 95% of probability to obtain a positive detection was 2623 target genes.

Supplementary tables

Table S1. 16S rRNA gene sequences used to perform oligonucleotides design.

Accession number	Description
AJ413954*	*Faecalibacterium prausnitzii*, strain ATCC 27768
X85022*	*F. prausnitzii*, strain ATCC 27766
AY305307*	Butyrate-producing bacterium M21/2
HQ457025*	*F. prausnitzii,* strain S4L/4
HQ457024*	*F. prausnitzii,* strain S3L/3
AJ270470*	Butyrate-producing bacterium L2-6
AJ270469*	Butyrate-producing bacterium A2-165
HQ457026*	*F. prausnitzii,* strain HTF-A
HQ457027*	*F. prausnitzii,* strain HTF-B
HQ457028*	*F. prausnitzii,* strain HTF-C
HQ457029*	*F. prausnitzii,* strain HTF-E
HQ457030*	*F. prausnitzii,* strain HTF-F
HQ457031*	*F. prausnitzii,* strain HTF-I
HQ457032*	*F. prausnitzii,* strain HTF-60C
HQ457033*	*F. prausnitzii,* strain HTF-75H
JN037415*	*F. prausnitzii,* strain L2-15
JN037416*	*F. prausnitzii,* strain L2-39
JN037417*	*F. prausnitzii,* strain L2-61
AM075671*	Uncultured bacterium partial 16S rRNA gene, isolate C0703§.
AM075691*	Uncultured bacterium partial 16S rRNA gene, isolate C0402§.
AM075696*	Uncultured bacterium partial 16S rRNA gene, isolate C0801§.
AM075683*	Uncultured bacterium partial 16S rRNA gene, isolate C1403§.
AM075730*	Uncultured bacterium partial 16S rRNA gene, isolate CD1902§.
AM075738*	Uncultured bacterium partial 16S rRNA gene, isolate UC0102§.
AY169429*	*Faecalibacterium prausnitzii* clone 1-84, partial sequence
AY169430*	*Faecalibacterium prausnitzii* clone 1-88, partial sequence
AY169427*	*Faecalibacterium prausnitzii* clone 1-79, partial sequence
AF132237*	Uncultured bacterium adhufec13, partial sequence§
AF132236*	Uncultured bacterium adhufec113, partial sequence§
AF132246*	Uncultured bacterium adhufec218, partial sequence§
AF132265*	Uncultured bacterium adhufec365, partial sequence§
AF153871*	Uncultured bacterium adhufec08.25, partial sequence§
X98011	*Anaerofilum agile*
X97852	*Anaerofilum pentosovorans*
AJ315980	*Anaerotruncus colihominis* DSM 17241.
AJ518869	*Subdoligranulum variabile* type strain BI 114^T
L09173	*Clostridium thermocellum* DSM 1237
L09177	*Clostridium cellulosi*
M59095	*Clostridium leptum*
AJ305238	*Clostridium leptum*; DSM 753^T
M59116	*Clostridium sporosphaeroides*
X66002	*Clostridium sporosphaeroides* DSM 1294
X81125	*Clostridium viride*
L34618	*Eubacterium desmolans*
L34625	*Eubacterium siraeum*
AY445600	*Ruminococcus albus,* strain 7, complete gene sequence
X85098	*Ruminococcus albus*
AY445594	*Ruminococcus albus,* strain 8 complete gene sequence
AY445592	*Ruminococcus albus,* strain B199 complete gene sequence
AY445596	*Ruminococcus albus,* strain KF1 complete gene sequence
AY445602	*Ruminococcus albus,* strain R013 complete gene sequence
X85099	*Ruminococcus bromii*
L76600	*Ruminococcus bromii*
X85100	*Ruminococcus callidus*

Accession number	Description
L76596	*Ruminococcus callidus*
X85097	*Ruminococcus flavefaciens*
AM915269	*Ruminococcus flavefaciens*, type strain C94^{T}=ATCC19208, partial sequence
AF030449	*Ruminococcus flavefaciens,* strain ATCC 49949, partial sequence
AY445599	*Ruminococcus flavefaciens,* strain B146, complete sequence
AY445597	*Ruminococcus flavefaciens,* strain FD1, complete sequence
AY445595	*Ruminococcus flavefaciens,* strain JM1, complete sequence
AY445593	*Ruminococcus flavefaciens,* strain C94, complete sequence
AY445603	*Ruminococcus flavefaciens,* strain LB4, complete sequence
AY445601	*Ruminococcus flavefaciens,* strain JF1, complete sequence
AY445598	*Ruminococcus flavefaciens,* strain R13e2, complete sequence

* Sequences used to obtain the *F. prausnitzii* 16S rRNA gene consensus sequence for oligonucleotides design
§ Sequences of the genus *Faecalibacterium*

Table S2. Growth conditions and source of the bacterial strains used in this study. The results obtained from the specificity tests are also included.

Type and origin of bacterial strains*		Growth (2)		Specificity test (3)	
Species/strains	Culture collection isolate (1)	Media	T(ºC)	Bacteria PCR	*F. prausnitzii* PCR
Firmicutes					
Faecalibacterium prausnitzii ATCC 27768^T	ATCC 27768	M2GSC	37	+	+
F. prausnitzii A2-165	DSM17677	M2GSC	37	+	+
F. prausnitzii M21/2	nd	M2GSC	37	+	+
F. prausnitzii L2-15	nd	M2GSC	37	+	+
F. prausnitzii L2-39	nd	M2GSC	37	+	+
F. prausnitzii L2-6	nd	M2GSC	37	+	+
F. prausnitzii L2-61	nd	M2GSC	37	+	+
F. prausnitzii S3L/3	nd	M2GSC	37	+	+
F. prausnitzii S4L/4	nd	M2GSC	37	+	+
Anaerofilum agile	DSM 4272	nc	nc	+	-
Eubacterium siraeum	DSMZ 15702	nc	nc	+	-
Eubacterium halii	DSMZ 17630	nc	nc	+	-
Clostridium viride	DSM 6836	nc	nc	+	-
Clostridium leptum	DSM 753	nc	nc	+	-
Ruminococcus albus	DSM 20455	nc	nc	+	-
Clostridium acetobutylicum	CECT 979	AN	37	+	-
Clostridium botulinum type E	CECT 4611	LiB	37	+	-
Bacillus cereus	NCTC 11145	AN	30	+	-
Bacillus megaterium	DSM 319	AN	30	+	-
Bacillus sp.	CECT 40	AN	30	+	-
Bacillus subtilis	NCTC 10400	AN	30	+	-
Bacillus subtilis sups. spizizwnii	CECT 482	AN	30	+	-
Listeria grayi	CECT931	BHI	37	+	-
Listeria innocua	CECT 910	BHI	37	+	-
Paenibacillus polymyxa	DSM 372	BHI	37	+	-
Staphylococcus aureus	ATCC 9144	AN	37	+	-
Staphylococcus epidermidis	CECT 231	AN	37	+	-
Enterococcus avium	CECT 968	BHI	37	+	-
Enterococcus columbae	CECT 4798	BHI	37	+	-
Enterococcus durans	CECT 411	BHI	37	+	-
Enterococcus faecalis	CECT 481	BHI	37	+	-
Enterococcus faecium	CECT 410	BHI	37	+	-
Enterococcus gallinarum	CECT 970	BHI	37	+	-
Enterococcus mundtii	CECT 972	BHI	37	+	-
Lactobacillus acidophilus	CECT 903	MRS	30	+	-
Lactococcus lactis	CECT 185	MRS	30	+	-
Streptococcus agalactiae	CECT 183	BHI	37	+	-
Streptococcus anginosus	CECT 948	BHI	37	+	-
Streptococcus equi subsp. *equi*	CECT 989	BHI	37	+	-
Streptococcus equinus	CECT 213	BHI	37	+	-
Streptococcus intermedius	CECT 803	BHI	37	+	-
Streptococcus mutans	CECT 479	BHI	37	+	-
Streptococcus oralis	CECT 907	BHI	37	+	-
Streptococcus pneumoniae	CECT 993	BHI	37	+	-
Streptococcus pyogenes	CECT 598	BHI	37	+	-
Streptococcus salivarus	CECT 805	BHI	37	+	-
Streptococcus sanguinis	CECT 480	BHI	37	+	-
Streptococcus sobrinus	CECT 4034	BHI	37	+	-
Streptococcus suis	CECT 958	BHI	37	+	-
Streptococcus thermophilus	CECT 986	BHI	37	+	-
Streptococcus uberis	CECT 994	BHI	37	+	-
Actinobacteria					
Corynebacterium bovis	DSM 20582	MRS	37	+	-

Type and origin of bacterial strains*		Growth [2]		Specificity test [3]	
Species/strains	Culture collection isolate [1]	Media	T(ºC)	Bacteria PCR	*F. prausnitzii* PCR
Kocuria rhizophila	DSM 348	AN	30	+	-
Micrococcus luteus	CECT 241	AN	30	+	-
Mycobacterium phlei	CECT 3009	BHI	37	+	-
Streptomyces griseus	DSM 40236	PDA	30	+	-
Bifidobacterium adolescentis	CECT 5781	AN	37	+	-
Bifidobacterium breve	CECT 4839	AN	37	+	-
Bacteroidetes					
Bacteroides fragilis	DSMZ 2151	nc	nc	+	-
Bacteroides uniformis	DSMZ 6597	nc	nc	+	-
Bacteroides vulgatus	DSMZ 1447	nc	nc	+	-
Proteobacteria					
Methylophilus methylotrophus	DSM 5691	CZ	30	+	-
Campylobacter jejuni	DSM 4688	BA	37	+	-
Citrobacter freundii	CECT 401	AN	30	+	-
Enterobacter aerogenes	CECT 684	AN	30	+	-
Enterobacter cloacae	CECT 194	AN	30	+	-
Enterobacter sakazakii	CECT 858	AN	30	+	-
Enterobacter sakazakii	ATCC 51329	AN	30	+	-
Enterobacter amnigenus (Sakazakii)	CECT 4078	AN	37	+	-
Enterobacter gergoviae (Sakazakii)	CECT 857	AN	37	+	-
Escherichia coli	CECT 100	AN	37	+	-
Escherichia coli	CECT 101	AN	37	+	-
Escherichia coli	CECT 105	AN	37	+	-
Escherichia coli	CECT 12242	AN	37	+	-
Escherichia coli	CECT 831	AN	37	+	-
Escherichia coli	CECT 4201	AN	37	+	-
Escherichia coli	CECT 4084	AN	37	+	-
Escherichia coli	CECT 405	AN	37	+	-
Escherichia coli	ATCC 10536	AN	37	+	-
Klebsiella pneumoniae ssp. pneumoniae	CECT 143	AN	37	+	-
Proteus mirabilis	CECT 170	AN	37	+	-
Salmonella LT2	CECT 878	AN	37	+	-
Salmonella TA98	CECT 880	AN	37	+	-
Serratia marcescens	CECT 846	AN	25	+	-
Shigella sonnei	CECT 457	AN	37	+	-
Pseudomonas aeruginosa	CECT 532	AN	30	+	-
Pseudomonas fluorescens	CECT 378	AN	30	+	-
Pseudomonas mendocina	CECT320	AN	30	+	-
Pseudomonas putida	CECT 324	AN	30	+	-

* Specificity test with human Xsomal DNA (Eurogentec, Belgium) was also performed

[1] ATCC: American Type Culture Collection (Manassas, VA, USA); CECT: Colección Española de Cultivos Tipo (Valencia, Spain); DSMZ: Deutche Sammlung von Mikroorganismen and Zellkulturen (Braunschweig, Germany), NCTC: National Collection of Type Cultures (London,UK), nd: not deposited (stocks held by the authors, Rowett Institute of nutrition and Health, Aberdeen, United Kingdom).

[2] nc: not cultured. BHI (Brain Heart Infusion Broth), AN (Nutrient Agar), BA (Blood Agar), MRS (Man, Rogosa and Sharpe medium), LiB (Liver Broth, CECT medium #15), CZ (Colby and Zathman medium, DSMZ medium #606), PDA (Potato Dextrose Agar), M2GSC (modified Med2 of Hobson, (5))

[3] The DNA was obtained from 1ml of bacterial culture at the stationary growth phase or for not cultured strains, the dried culture directly obtained from the culture type collection was rehydrated with the appropriate buffer for DNA extraction and used to DNA purification. Up to 10ng of DNA were used to perform the test

Table S3. Different unique 16S rRNA gene sequences found in this study as identified by Mothur. Frequency of detection of each unique sequence has been specified by group of patients.

	Number of sequences							Nearest *F. prausnitzii* isolate			Nearest sequence		
Sequence unique	H	IBS	UC	CD	CRC	Total	OTUrep*	Strain	Similarity (%)	Accession Number	Similarity (%)	Accession Number	Description
OTU100_001	0	1	8	13	9	31	KP005701	HTF-I	99	HQ457031.1	99	JQ941003.1	Uncultured bacterium clone 2118007-1-71
OTU100_002	4	1	4	5	4	18	KP005691	HTF-I	99	HQ457031.1	100	JQ940601.1	Uncultured bacterium clone 218007-192
OTU100_003	0	1	2	9	4	16	KP005681	HTF-I	99	HQ457031.1	99	JQ940533.1	Uncultured bacterium clone 218004-1-70
OTU100_004	1	1	1	7	3	13	KP005705	M21/2	99	AY305307.1	100	KF101872.1	Uncultured bacterium clone ncd2742a12c1
OTU100_005	1	0	3	5	4	13	KP005704	L2-6	99	AJ270470.2	99	KF071154.1	Uncultured bacterium clone ncd245d04c1
OTU100_006	5	2	2	1	1	11	KP005667	S3L/3	99	HQ457024.1	100	JQ940877.1	Uncultured bacterium clone 218004-1-94
OTU100_007	4	0	2	4	1	11	KP005675	HTF-I	99	HQ457031.1	99	HQ813966.1	Uncultured organism clone ELU0171-T442-S-NIPCRAMgANa_000442
OTU100_008	1	3	2	3	1	10	KP005706	L2-6	99	AJ270470.2	100	JQ189863.1	Uncultured bacterium clone BD16778
OTU100_009	1	0	0	5	2	8	KP005720	M21/2	100	AY305307.1	100	JQ941056.1	Uncultured bacterium clone 218004-1-120
OTU100_010	2	0	1	2	2	7	KP005684	S4L/4	98	HQ457025.1	100	KF843160.1	Uncultured bacterium clone SG_B476
OTU100_011	0	0	3	1	1	5	KP005623	A2-165	99	AJ270469.2	99	FJ510187.1	Uncultured bacterium clone 16slp75-10g08.p1k
OTU100_012	1	0	2	0	1	4	KP005562	M21/2	99	AY305307.1	99	JQ941082.1	Uncultured bacterium clone 218004-1-61
OTU100_013	1	1	2	0	0	4	KP005574	S3L/3	99	HQ457024.1	99	HQ777603.1	Uncultured organism clone ELU0082-T384-S-NI_000369
OTU100_014	0	0	0	4	0	4	KP005712	M21/2	99	AY305307.1	100	HQ806136.1	Uncultured organism clone ELU0156-T284-S-NIPCRAMgANa_000654
OTU100_015	0	0	2	0	1	3	KP005678	HTF-I	99	HQ457031.1	99	JQ941003.1	Uncultured bacterium clone 2118007-1-71
OTU100_016	0	0	0	2	1	3	KP005595	HTF-I	99	HQ457031.1	99	HQ813966.1	Uncultured organism clone ELU0171-T442-S-NIPCRAMgANa_000442
OTU100_017	1	0	1	0	0	2	KP005490	HTF-I	99	HQ457031.1	99	JQ941003.1	Uncultured bacterium clone 2118007-1-71
OTU100_018	0	0	0	0	2	2	KP005661	HTF-I	99	HQ457031.1	99	JQ940601.1	Uncultured bacterium clone 218007-192
OTU100_019	0	0	0	2	0	2	KP005592	L2-6	97	AJ270470.2	99	KF088388.1	Uncultured bacterium clone nck231a09c1
OTU100_020	1	1	0	0	0	2	KP005472	M21/2	99	AY305307.1	99	JQ190737.1	Uncultured bacterium clone BD09374
OTU100_021	0	0	0	1	0	1	KP005643	HTF-I	99	HQ457031.1	99	JQ941003.1	Uncultured bacterium clone 2118007-1-71
OTU100_022	0	0	0	1	0	1	KP005625	L2-6	99	AJ270470.2	99	HQ791201.1	Uncultured organism clone ELU0123-T309-S-NI_000213
OTU100_023	0	0	0	1	0	1	KP005628	S3L/3	98	HQ457024.1	99	JQ190472.1	Uncultured bacterium clone BD17036
OTU100_024	0	0	0	1	0	1	KP005629	L2-6	99	AJ270470.2	99	JQ189863.1	Uncultured bacterium clone BD16778

	Number of sequences							Nearest *F. prausnitzii* isolate			Nearest sequence		
Sequence unique	H	IBS	UC	CD	CRC	Total	OTUrep*	Strain	Similarity (%)	Accession Number	Similarity (%)	Accession Number	Description
OTU100_025	0	0	0	1	0	1	KP005631	L2-6	99	AJ270470.2	99	JQ189863.1	Uncultured bacterium clone BD16778
OTU100_026	0	0	0	1	0	1	KP005633	S3L/3	98	HQ457024.1	99	JQ190472.1	Uncultured bacterium clone BD17036
OTU100_027	0	0	1	0	0	1	KP005559	A2-165	99	AJ270469.2	99	HQ777603.1	Uncultured organism clone ELU0082-T384-S-NI_000369
OTU100_028	0	0	0	1	0	1	KP005638	HTF-I	99	HQ457031.1	99	JQ941003.1	Uncultured bacterium clone 2118007-1-71
OTU100_029	0	0	0	1	0	1	KP005642	L2-6	99	AJ270470.2	99	JQ189863.1	Uncultured bacterium clone BD16778
OTU100_030	0	0	0	1	0	1	KP005620	M21/2	99	AY305307.1	99	DQ326121.1	Uncultured bacterium clone BB84
OTU100_031	0	0	0	1	0	1	KP005644	M21/2	99	AY305307.1	99	GQ896754.1	Uncultured bacterium clone C2-7
OTU100_032	0	0	0	0	1	1	KP005648	L2-6	99	AJ270470.2	99	FP083339.1	16S rDNA sequence amplified from human fecal sample
OTU100_033	0	0	0	0	1	1	KP005650	S3L/3	98	HQ457024.1	99	JQ190472.1	Uncultured bacterium clone BD17036
OTU100_034	0	1	0	0	0	1	KP005556	A2-165	99	AJ270469.2	99	FJ510187.1	Uncultured bacterium clone 16slp75-10g08.p1k
OTU100_035	0	0	0	0	1	1	KP005652	M21/2	99	AY305307.1	99	JQ941082.1	Uncultured bacterium clone 218004-1-61
OTU100_036	0	1	0	0	0	1	KP005552	L2-6	99	AJ270470.2	99	FP083606.1	16S rDNA sequence amplified from human fecal sample
OTU100_037	0	1	0	0	0	1	KP005551	A2-165	97	AJ270469.2	98	JQ189747.1	Uncultured bacterium clone BD16481
OTU100_038	0	0	1	0	0	1	KP005577	HTF-I	99	HQ457031.1	99	HQ813966.1	Uncultured organism clone ELU0171-T442-S-NIPCRAMgANa_000442
OTU100_039	0	0	0	1	0	1	KP005584	L2-6	99	AJ270470.2	99	KF071154.1	Uncultured bacterium clone ncd245d04c1
OTU100_040	0	0	0	1	0	1	KP005583	HTF-I	98	HQ457031.1	99	EF404739.1	Uncultured bacterium clone SJTU_C_15_63
OTU100_041	0	0	1	0	0	1	KP005578	S3L/3	99	HQ457024.1	99	HQ777603.1	Uncultured organism clone ELU0082-T384-S-NI_000369
OTU100_042	0	0	0	1	0	1	KP005594	HTF-I	99	HQ457031.1	99	DQ802202.1	Uncultured bacterium clone RL241_aaj04d07
OTU100_043	0	0	0	1	0	1	KP005599	HTF-I	99	HQ457031.1	99	DQ802202.1	Uncultured bacterium clone RL241_aaj04d07
OTU100_044	0	0	0	1	0	1	KP005601	L2-6	99	AJ270470.2	99	KF071154.1	Uncultured bacterium clone ncd245d04c1
OTU100_045	0	0	0	1	0	1	KP005606	M21/2	99	AY305307.1	99	KF101872.1	Uncultured bacterium clone ncd2742a12c1
OTU100_046	0	0	0	0	1	1	KP005676	L2-6	97	AJ270470.2	98	KF088388.1	Uncultured bacterium clone nck231a09c1
OTU100_047	0	0	0	1	0	1	KP005593	HTF-I	99	HQ457031.1	99	JQ941003.1	Uncultured bacterium clone 2118007-1-71
OTU100_048	0	0	1	0	0	1	KP005576	HTF-I	98	HQ457031.1	99	FJ504122.1	Uncultured bacterium clone 16slp55-04a02.p1k
OTU100_049	0	0	0	1	0	1	KP005610	HTF-I	99	HQ457031.1	99	DQ802202.1	Uncultured bacterium clone RL241_aaj04d07
OTU100_050	0	0	0	1	0	1	KP005615	HTF-I	99	HQ457031.1	99	FP083067.1	16S rDNA sequence amplified from human fecal sample
OTU100_051	0	0	0	1	0	1	KP005616	L2-6	99	AJ270470.2	99	EF405391.1	Uncultured bacterium clone SJTU_G_01_17

	Number of sequences							Nearest *F. prausnitzii* isolate			Nearest sequence		
Sequence unique	H	IBS	UC	CD	CRC	Total	OTUrep*	Strain	Similarity (%)	Accession Number	Similarity (%)	Accession Number	Description
OTU100_052	0	0	1	0	0	1	KP005571	S3L/3	99	HQ457024.1	99	HQ777603.1	Uncultured organism clone ELU0082-T384-S-NI_000369
OTU100_053	0	0	1	0	0	1	KP005566	L2-6	98	AJ270470.2	98	JQ941003.1	Uncultured bacterium clone 2118007-1-71
OTU100_054	0	0	0	1	0	1	KP005618	HTF-I	99	HQ457031.1	99	JQ941003.1	Uncultured bacterium clone 2118007-1-71
OTU100_055	1	0	0	0	0	1	KP005731	S4L/4	98	HQ457025.1	99	JQ189747.1	Uncultured bacterium clone BD16481
OTU100_056	0	0	0	1	0	1	KP005716	M21/2	99	AY305307.1	99	JQ941056.1	Uncultured bacterium clone 218004-1-120
OTU100_057	0	0	0	1	0	1	KP005717	L2-6	99	AJ270470.2	99	KF071154.1	Uncultured bacterium clone ncd245d04c1
OTU100_058	0	0	0	1	0	1	KP005718	HTF-I	98	HQ457031.1	99	HQ813966.1	Uncultured organism clone ELU0171-T442-S-NIPCRAMgANa_000442
OTU100_059	0	0	0	1	0	1	KP005719	S4L/4	99	HQ457025.1	99	JQ940877.1	Uncultured bacterium clone 218004-1-94
OTU100_060	0	0	0	1	0	1	KP005724	HTF-I	99	HQ457031.1	99	FJ504122.1	Uncultured bacterium clone 16slp55-04a02.p1k
OTU100_061	0	0	0	1	0	1	KP005726	L2-6	99	AJ270470.2	99	KF071154.1	Uncultured bacterium clone ncd245d04c1
OTU100_062	1	0	0	0	0	1	KP005727	M21/2	95§	AY305307.1	95§	JQ189840.1	Uncultured bacterium clone BD16180
OTU100_063	1	0	0	0	0	1	KP005728	M21/2	99	AY305307.1	99	JQ941082.1	Uncultured bacterium clone 218004-1-61
OTU100_064	1	0	0	0	0	1	KP005729	S4L/4	99	HQ457025.1	99	FJ683640.1	Uncultured bacterium clone 8-7I8
OTU100_065	0	0	0	1	0	1	KP005715	HTF-I	98	HQ457031.1	99	JQ940656.1	Uncultured bacterium clone 218004-1-73
OTU100_066	1	0	0	0	0	1	KP005732	HTF-I	99	HQ457031.1	99	HQ813966.1	Uncultured organism clone ELU0171-T442-S-NIPCRAMgANa_000442
OTU100_067	1	0	0	0	0	1	KP005733	S4L/4	99	HQ457025.1	99	JQ190472.1	Uncultured bacterium clone BD17036
OTU100_068	1	0	0	0	0	1	KP005734	HTF-I	99	HQ457031.1	99	EF404739.1	Uncultured bacterium clone SJTU_C_15_63
OTU100_069	1	0	0	0	0	1	KP005735	M21/2	99	AY305307.1	99	JQ941082.1	Uncultured bacterium clone 218004-1-61
OTU100_070	0	0	0	0	1	1	KP005737	HTF-I	98	HQ457031.1	99	HQ778999.1	Uncultured organism clone ELU0086-T395-S-NI_000183
OTU100_071	0	0	0	0	1	1	KP005738	HTF-I	98	HQ457031.1	99	FP083067.1	16S rDNA sequence amplified from human fecal sample
OTU100_072	0	0	0	0	1	1	KP005739	HTF-I	98	HQ457031.1	99	JQ940601.1	Uncultured bacterium clone 218007-192
OTU100_073	0	0	0	1	0	1	KP005741	M21/2	99	AY305307.1	100	HQ806136.1	Uncultured organism clone ELU0156-T284-S-NIPCRAMgANa_000654
OTU100_074	0	0	0	0	1	1	KP005689	A2-165	99	AJ270469.2	100	HQ810293.1	Uncultured organism clone ELU0164-T379-S-NIPCRAMgANa_000241
OTU100_075	0	0	0	0	1	1	KP005663	HTF-I	99	HQ457031.1	99	HQ813966.1	Uncultured organism clone ELU0171-T442-S-NIPCRAMgANa_000442
OTU100_076	0	0	0	0	1	1	KP005666	HTF-I	98	HQ457031.1	99	JQ940656.1	Uncultured bacterium clone 218004-1-73
OTU100_077	0	0	0	0	1	1	KP005668	M21/2	99	AY305307.1	99	JQ941056.1	Uncultured bacterium clone 218004-1-120
OTU100_078	0	0	0	0	1	1	KP005669	M21/2	99	AY305307.1	99	JQ941056.1	Uncultured bacterium clone 218004-1-120

	Number of sequences							Nearest *F. prausnitzii* isolate			Nearest sequence		
Sequence unique	H	IBS	UC	CD	CRC	Total	OTUrep*	Strain	Similarity (%)	Accession Number	Similarity (%)	Accession Number	Description
OTU100_079	1	0	0	0	0	1	KP005459	L2-6	99	AJ270470.2	99	KF071154.1	Uncultured bacterium clone ncd245d04c1
OTU100_080	0	0	0	0	1	1	KP005677	M21/2	99	AY305307.1	99	JQ189840.1	Uncultured bacterium clone BD16180
OTU100_081	0	0	0	0	1	1	KP005683	HTF-I	98	HQ457031.1	98	HQ813966.1	Uncultured organism clone ELU0171-T442-S-NIPCRAMgANa_000442
OTU100_082	0	0	0	0	1	1	KP005688	A2-165	99	AJ270469.2	99	JQ189747.1	Uncultured bacterium clone BD16481
OTU100_083	0	0	0	0	1	1	KP005662	S3L/3	99	HQ457024.1	99	HQ777603.1	Uncultured organism clone ELU0082-T384-S-NI_000369
OTU100_084	0	0	0	0	1	1	KP005690	A2-165	99	AJ270469.2	99	HQ820492.1	Uncultured organism clone ELU0181-T360-S-NIPCRAMgANa_000532
OTU100_085	0	0	0	0	1	1	KP005693	A2-165	99	AJ270469.2	99	JQ189747.1	Uncultured bacterium clone BD16481
OTU100_086	0	0	0	0	1	1	KP005697	L2-6	99	AJ270470.2	99	KF071154.1	Uncultured bacterium clone ncd245d04c1
OTU100_087	0	0	0	1	0	1	KP005699	A2-165	99	AJ270469.2	99	HQ764137.1	Uncultured organism clone ELU0055-T366-S-NIPCRAMgANa_000271
OTU100_088	0	0	1	0	0	1	KP005702	M21/2	99	AY305307.1	99	KF101872.1	Uncultured bacterium clone ncd2742a12c1
OTU100_089	0	0	1	0	0	1	KP005708	L2-6	99	AJ270470.2	99	KF071154.1	Uncultured bacterium clone ncd245d04c1
OTU100_090	0	0	0	1	0	1	KP005709	A2-165	99	AJ270469.2	99	FJ510187.1	Uncultured bacterium clone 16slp75-10g08.p1k
OTU100_091	0	0	0	1	0	1	KP005710	M21/2	98	AY305307.1	99	FJ683640.1	Uncultured bacterium clone 8-7I8
OTU100_092	1	0	0	0	0	1	KP005533	L2-6	99	AJ270470.2	99	KF071154.1	Uncultured bacterium clone ncd245d04c1
OTU100_093	1	0	0	0	0	1	KP005539	S3L/3	99	HQ457024.1	99	HQ777603.1	Uncultured organism clone ELU0082-T384-S-NI_000369
OTU100_094	1	0	0	0	0	1	KP005509	M21/2	98	AY305307.1	99	JQ941082.1	Uncultured bacterium clone 218004-1-61
OTU100_095	0	0	1	0	0	1	KP005483	HTF-I	99	HQ457031.1	99	HQ813966.1	Uncultured organism clone ELU0171-T442-S-NIPCRAMgANa_000442
OTU100_096	1	0	0	0	0	1	KP005516	L2-6	99	AJ270470.2	99	FP083339.1	16S rDNA sequence amplified from human fecal sample
OTU100_097	1	0	0	0	0	1	KP005508	S4L/4	99	HQ457025.1	99	JQ940877.1	Uncultured bacterium clone 218004-1-94
OTU100_098	0	0	1	0	0	1	KP005486	M21/2	98	AY305307.1	99	HM479020.1	Uncultured organism clone UUAV8AF101
OTU100_099	0	0	1	0	0	1	KP005487	HTF-I	99	HQ457031.1	99	JQ941003.1	Uncultured bacterium clone 2118007-1-71
OTU100_100	1	0	0	0	0	1	KP005507	S3L/3	99	HQ457024.1	100	HM286763.1	Uncultured bacterium clone ncd634b09c1
OTU100_101	1	0	0	0	0	1	KP005506	HTF-I	99	HQ457031.1	99	JQ940601.1	Uncultured bacterium clone 218007-192
OTU100_102	1	0	0	0	0	1	KP005519	S3L/3	99	HQ457024.1	99	JQ940877.1	Uncultured bacterium clone 218004-1-94
OTU100_103	1	0	0	0	0	1	KP005535	M21/2	99	AY305307.1	99	JQ941082.1	Uncultured bacterium clone 218004-1-61
OTU100_104	1	0	0	0	0	1	KP005458	M21/2	99	AY305307.1	99	JQ189471.1	Uncultured bacterium clone BD15679
OTU100_105	1	0	0	0	0	1	KP005532	HTF-I	99	HQ457031.1	99	HQ813966.1	Uncultured organism clone ELU0171-T442-S-NIPCRAMgANa_000442

	Number of sequences							Nearest *F. prausnitzii* isolate			Nearest sequence		
Sequence unique	H	IBS	UC	CD	CRC	Total	OTUrep*	Strain	Similarity (%)	Accession Number	Similarity (%)	Accession Number	Description
OTU100_106	1	0	0	0	0	1	KP005522	S4L/4	98	HQ457025.1	99	KF843160.1	Uncultured bacterium clone SG_B476
OTU100_107	1	0	0	0	0	1	KP005523	S4L/4	99	HQ457025.1	99	HQ777603.1	Uncultured organism clone ELU0082-T384-S-NI_000369
OTU100_108	1	0	0	0	0	1	KP005524	S3L/3	99	HQ457024.1	99	HQ777603.1	Uncultured organism clone ELU0082-T384-S-NI_000369
OTU100_109	0	0	1	0	0	1	KP005497	M21/2	99	AY305307.1	99	JQ941056.1	Uncultured bacterium clone 218004-1-120
OTU100_110	0	0	0	1	0	1	KP005498	L2-6	99	AJ270470.2	99	JQ189863.1	Uncultured bacterium clone BD16778
OTU100_111	1	0	0	0	0	1	KP005529	L2-6	98	AJ270470.2	98	EU768088.1	Uncultured bacterium clone C4_553 16S
OTU100_112	0	0	0	1	0	1	KP005499	HTF-I	98	HQ457031.1	99	HQ778999.1	Uncultured organism clone ELU0086-T395-S-NI_000183
OTU100_113	1	0	0	0	0	1	KP005525	S3L/3	99	HQ457024.1	99	FP080037.1	16S rDNA sequence amplified from human fecal sample
OTU100_114	1	0	0	0	0	1	KP005526	HTF-I	98	HQ457031.1	99	JQ940601.1	Uncultured bacterium clone 218007-192
OTU100_115	1	0	0	0	0	1	KP005527	S3L/3	99	HQ457024.1	99	JQ190987.1	Uncultured bacterium clone BD09145
OTU100_116	1	0	0	0	0	1	KP005463	M21/2	99	AY305307.1	99	JQ186494.1	Uncultured bacterium clone BD07222
OTU100_117	1	0	0	0	0	1	KP005466	S4L/4	99	HQ457025.1	99	FJ683640.1	Uncultured bacterium clone 8-7I8
OTU100_118	0	1	0	0	0	1	KP005544	A2-165	99	AJ270469.2	99	JQ189747.1	Uncultured bacterium clone BD16481
OTU100_119	1	0	0	0	0	1	KP005467	S4L/4	98	HQ457025.1	99	JQ190472.1	Uncultured bacterium clone BD17036
OTU100_120	1	0	0	0	0	1	KP005465	M21/2	99	AY305307.1	99	JQ189840.1	Uncultured bacterium clone BD16180
OTU100_121	1	0	0	0	0	1	KP005468	M21/2	99	AY305307.1	99	JQ941056.1	Uncultured bacterium clone 218004-1-120
OTU100_122	1	0	0	0	0	1	KP005469	M21/2	99	AY305307.1	99	JQ190737.1	Uncultured bacterium clone BD09374
OTU100_123	0	1	0	0	0	1	KP005545	M21/2	99	AY305307.1	99	KF101872.1	Uncultured bacterium clone ncd2742a12c1
OTU100_124	1	0	0	0	0	1	KP005510	L2-6	99	AJ270470.2	99	KF071154.1	Uncultured bacterium clone ncd245d04c1
OTU100_125	1	0	0	0	0	1	KP005470	M21/2	99	AY305307.1	99	JQ189840.1	Uncultured bacterium clone BD16180
OTU100_126	1	0	0	0	0	1	KP005471	M21/2	99	AY305307.1	99	JQ190737.1	Uncultured bacterium clone BD09374
OTU100_127	1	0	0	0	0	1	KP005462	S4L/4	98	HQ457025.1	99	FP080831.1	16S rDNA sequence amplified from human fecal sample
OTU100_128	1	0	0	0	0	1	KP005477	HTF-I	99	HQ457031.1	99	FP083067.1	16S rDNA sequence amplified from human fecal sample
OTU100_129	1	0	0	0	0	1	KP005460	S3L/3	97	HQ457024.1	98	JQ940828.1	Uncultured bacterium clone 218002-353
OTU100_130	0	0	1	0	0	1	KP005478	S4L/4	99	HQ457025.1	99	FJ683640.1	Uncultured bacterium clone 8-7I8
OTU100_131	0	1	0	0	0	1	KP005542	S4L/4	99	HQ457025.1	99	JQ186908.1	Uncultured bacterium clone BD08135
OTU100_132	1	0	0	0	0	1	KP005473	HTF-I	98	HQ457031.1	99	HQ813984.1	Uncultured organism clone ELU0171-T442-S-NIPCRAMgANa_00046ᴜ

	Number of sequences							Nearest *F. prausnitzii* isolate			Nearest sequence		
Sequence unique	H	IBS	UC	CD	CRC	Total	OTUrep*	Strain	Similarity (%)	Accession Number	Similarity (%)	Accession Number	Description
OTU100_133	1	0	0	0	0	1	KP005514	M21/2	99	AY305307.1	99	KF101872.1	Uncultured bacterium clone ncd2742a12c1
OTU100_134	1	0	0	0	0	1	KP005513	L2-6	98	AJ270470.2	99	HQ813666.1	Uncultured organism clone ELU0171-T442-S-NIPCRAMgANa_000142
OTU100_135	1	0	0	0	0	1	KP005461	L2-6	98	AJ270470.2	98	FP077644.1	16S rDNA sequence amplified from human fecal sample
Total sequences	66	17	48	97	56	284							

* Representative sequence when several sequences recovered from different subjects were identical
§ Novel phylotype with identity ≤95% to any NCBI/EMBL sequence

Table S4. Different phylotypes found as calculated by Mothur with the farthest neighbour method using a cut-off of 99% similarity of the 16S rRNA gene sequence. Frequency of detection of each OTU has been specified by group of patients.

OTU at 99% similarity	Number of sequences						OTUrep*	Nearest *F. prausnitzii* isolate			Nearest sequence		
	H	IBS	UC	CD	CRC	Total		Strain	Similarity (%)	Accession Number	Similarity (%)	Accession Number	Description
OTU99_1	16	3	22	41	22	104	KP005641	HTF-I	99	HQ457031.1	100	JQ941003.1	Uncultured bacterium clone 2118007-1-71
OTU99_2	7	3	5	19	9	43	KP005544	M21/2	99	AY305307.1	99	JQ189840.1	Uncultured bacterium clone BD16180
OTU99_3	6	4	6	18	7	41	KP005466	L2-6	99	AJ270470.2	99	KF071154.1	Uncultured bacterium clone ncd245d04c1
OTU99_4	14	4	10	3	5	36	KP005663	S3L/3	99	HQ457024.1	99	HQ777603.1	Uncultured organism clone ELU0082-T384-S-NI_000369
OTU99_5	3	0	1	2	2	8	KP005728	S4L/4	98	HQ457025.1	99	KF843160.1	Uncultured bacterium clone SG_B476
OTU99_6	2	0	0	1	1	4	KP005592	M21/2	99	AY305307.1	99	JQ190737.1	Uncultured bacterium clone BD09374
OTU99_7	0	0	0	2	2	4	KP005509	HTF-I	98	HQ457031.1	99	HQ778999.1	Uncultured organism clone ELU0086-T395-S-NI_000183
OTU99_8	0	0	0	2	1	3	KP005514	S3L/3	98	HQ457024.1	99	JQ190472.1	Uncultured bacterium clone BD17036
OTU99_9	3	0	0	0	0	3	KP005551	S4L/4	99	HQ457025.1	99	FJ683640.1	Uncultured bacterium clone 8-7I8
OTU99_10	0	1	0	0	2	3	KP005620	A2-165	99	AJ270469.2	99	JQ189747.1	Uncultured bacterium clone BD16481
OTU99_11	1	1	0	0	0	2	KP005710	S4L/4	99	HQ457025.1	99	FJ683640.1	Uncultured bacterium clone 8-7I8
OTU99_12	0	0	0	1	1	2	KP005465	HTF-I	99	HQ457031.1	99	HQ813966.1	Uncultured organism clone ELU0171-T442-S-NIPCRAMgANa_000442
OTU99_13	2	0	0	0	0	2	KP005583	M21/2	99	AY305307.1	99	JQ941082.1	Uncultured bacterium clone 218004-1-61
OTU99_14	0	0	0	2	0	2	KP005699	L2-6	97	AJ270470.2	99	KF088388.1	Uncultured bacterium clone nck231a09c1
OTU99_15	2	0	0	0	0	2	KP005513	M21/2	98	AY305307.1	99	JQ941082.1	Uncultured bacterium clone 218004-1-61
OTU99_16	1	0	0	0	0	1	KP005529	M21/2	99	AY305307.1	99	KF101872.1	Uncultured bacterium clone ncd2742a12c1
OTU99_17	0	1	0	0	0	1	KP005576	A2-165	97	AJ270469.2	98	JQ189747.1	Uncultured bacterium clone BD16481
OTU99_18	0	0	0	1	0	1	KP005644	M21/2	99	AY305307.1	99	DQ326121.1	Uncultured bacterium clone BB84
OTU99_19	0	0	0	1	0	1	KP005676	M21/2	98	AY305307.1	99	FJ683640.1	Uncultured bacterium clone 8-7I8
OTU99_20	0	0	0	1	0	1	KP005566	HTF-I	98	HQ457031.1	99	EF404739.1	Uncultured bacterium clone SJTU_C_15_63
OTU99_21	0	0	0	1	0	1	KP005683	A2-165	99	AJ270469.2	99	HQ764137.1	Uncultured organism clone ELU0055-T366-S-NIPCRAMgANa_000271
OTU99_22	1	0	0	0	0	1	KP005727	L2-6	98	AJ270470.2	99	HQ813666.1	Uncultured organism clone ELU0171-T442-S-NIPCRAMgANa_000142
OTU99_23	1	0	0	0	0	1	KP005694	L2-6	98	AJ270470.2	98	EU768088.1	Uncultured bacterium clone C4_553 16S
OTU99_24	0	0	1	0	0	1	KP005463	HTF-I	98	HQ457031.1	99	FJ504122.1	Uncultured bacterium clone 16slp55-04a02.p1k
OTU99_25	0	0	0	1	0	1	KP005478	M21/2	99	AY305307.1	99	GQ896754.1	Uncultured bacterium clone C2-7
OTU99_26	0	0	0	0	1	1	KP005462	L2-6	97	AJ270470.2	98	KF088388.1	Uncultured bacterium clone nck231a09c1
OTU99_27	0	0	1	0	0	1	KP005461	L2-6	98	AJ270470.2	98	JQ941003.1	Uncultured bacterium clone 2118007-1-71
OTU99_28	0	0	0	0	1	1	KP005739	HTF-I	98	HQ457031.1	98	HQ813966.1	Uncultured organism clone ELU0171-T442-S-NIPCRAMgANa_00044
OTU99_29	1	0	0	0	0	1	KP005460	M21/2	95§	AY305307.1	95§	JQ189840.1	Uncultured bacterium clone BD16180
OTU99_30	1	0	0	0	0	1	KP005738	M21/2	99	AY305307.1	99	JQ186494.1	Uncultured bacterium clone BD07222

OTU at 99% similarity	Number of sequences						OTUrep*	Nearest *F. prausnitzii* isolate			Nearest sequence		
	H	IBS	UC	CD	CRC	Total		Strain	Similarity (%)	Accession Number	Similarity (%)	Accession Number	Description
OTU99_31	0	0	1	0	0	1	KP005467	S4L/4	99	HQ457025.1	99	FJ683640.1	Uncultured bacterium clone 8-7I8
OTU99_32	1	0	0	0	0	1	KP005724	S4L/4	98	HQ457025.1	99	FP080831.1	16S rDNA sequence amplified from human fecal sample
OTU99_33	1	0	0	0	0	1	KP005486	L2-6	98	AJ270470.2	98	FP077644.1	16S rDNA sequence amplified from human fecal sample
OTU99_34	0	0	0	0	1	1	KP005574	HTF-I	98	HQ457031.1	99	JQ940601.1	Uncultured bacterium clone 218007-192
OTU99_35	1	0	0	0	0	1	KP005458	S3L/3	97	HQ457024.1	98	JQ940828.1	Uncultured bacterium clone 218002-353
OTU99_36	0	0	0	0	1	1	KP005504	HTF-I	98	HQ457031.1	99	FP083067.1	16S rDNA sequence amplified from human fecal sample
OTU99_37	1	0	0	0	0	1	KP005471	S4L/4	98	HQ457025.1	99	JQ190472.1	Uncultured bacterium clone BD17036
OTU99_38	0	0	0	1	0	1	KP005499	HTF-I	99	HQ457031.1	99	FJ504122.1	Uncultured bacterium clone 16slp55-04a02.p1k
OTU99_39	0	0	1	0	0	1	KP005650	M21/2	98	AY305307.1	99	HM479020.1	Uncultured organism clone UUAV8AF101
OTU99_40	1	0	0	0	0	1	KP005729	M21/2	99	AY305307.1	99	JQ189471.1	Uncultured bacterium clone BD15679
Total sequences	66	17	48	97	56	284							

* Representative sequence of this operational taxonomic unit (OTU)
§ Novel phylotype with identity ≤95% to any NCBI/EMBL sequence

Table S5. Different phylogroups found as calculated by Mothur with the farthest neighbour method using a cut-off of 97% similarity of the 16S rRNA gene sequence. Frequency of detection of each OTU has been specified by group of patients.

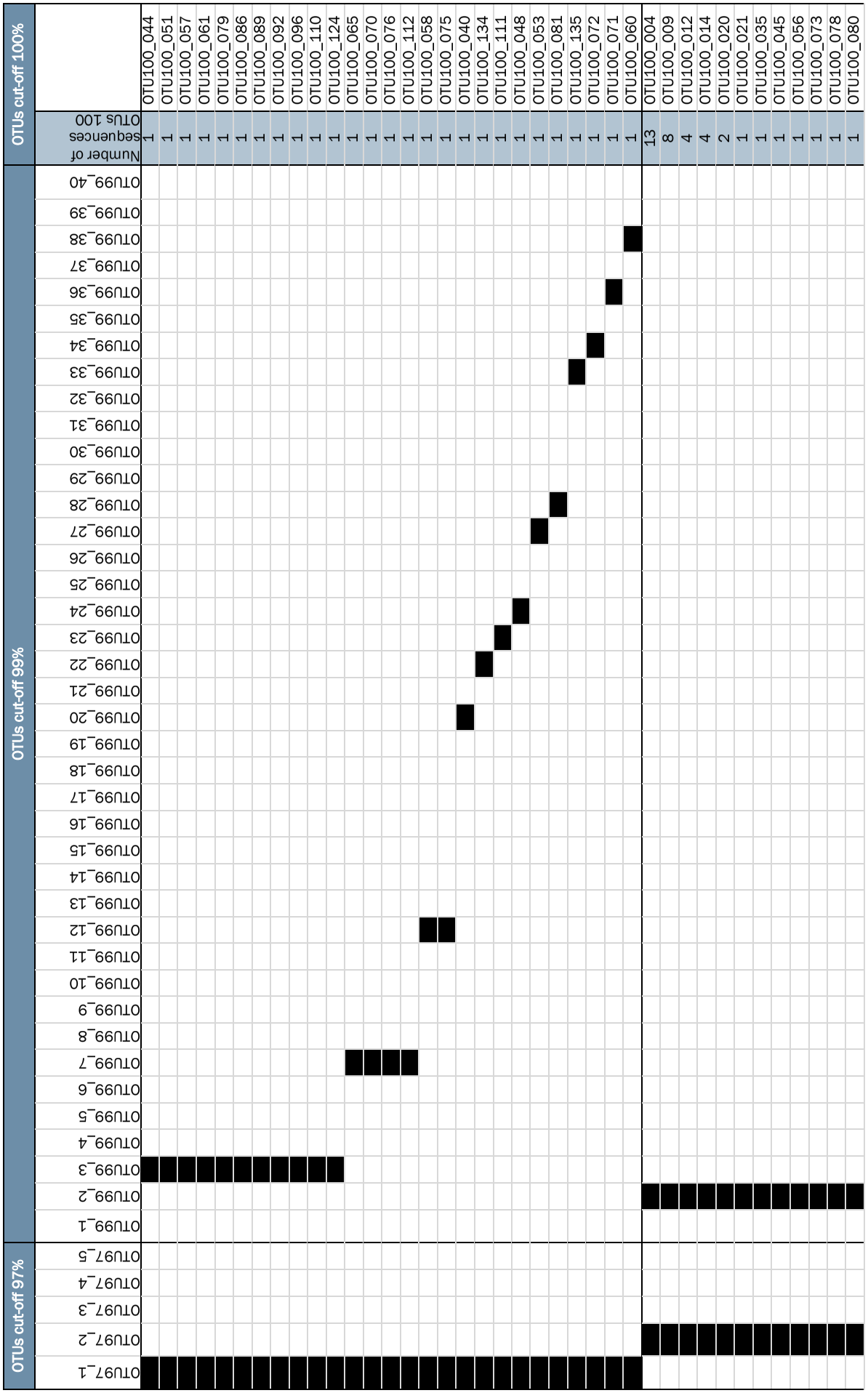
OTU at 97% similarity	Number of sequences						OTUrep*	Nearest *F. prausnitzii* isolate			Nearest sequence		
	H	IBS	UC	CD	CRC	Total		Strain	Similarity (%)	Accession Number	Similarity (%)	Accession Number	Description
OTU97_1	25	7	30	64	35	161	KP005568	HTF-I	99	HQ457031.1	99	JQ941003.1	Uncultured bacterium clone 2118007-1-71
OTU97_2	40	9	18	31	20	118	KP005574	S3L/3	99	HQ457024.1	99	HQ777603.1	Uncultured organism clone ELU0082-T384-S-NI_000369
OTU97_3	0	0	0	2	1	3	KP005592	L2/6	97	AJ270470.2	99	KF088388.1	Uncultured bacterium clone nck231a09c1
OTU97_4	0	1	0	0	0	1	KP005551	A2-165	97	AJ270469.2	98	JQ189747.1	Uncultured bacterium clone BD16481
OTU97_5	1	0	0	0	0	1	KP005727	M21/1	95§	AY305307.1	95§	JQ189840.1	Uncultured bacterium clone BD16180
Total sequences	66	17	48	97	56	284							

* Representative sequence of this operational taxonomic unit (OTU)

§ Novel phylotype with identity ≤95% to any NCBI/EMBL sequence is indicated

Table S6. Matrix showing the equivalence of OTU100 (unique sequences) with OTU99 (99% similarity of 16S rRNA gene sequence) and OTU97 (97% similarity of 16S rRNA gene sequence). Those OTU equivalent are shadowed in black: i.e. OTU100_1 corresponds to OTU99_1 and OTU97_1. Number of sequences included by each OTU are indicated.

OTUs cut-off 97%					OTUs cut-off 99%																																	OTUs cut-off 100%	
OTU97_1	OTU97_2	OTU97_3	OTU97_4	OTU97_5	OTU99_1	OTU99_2	OTU99_3	OTU99_4	OTU99_5	OTU99_6	OTU99_7	OTU99_8	OTU99_9	OTU99_10	OTU99_11	OTU99_12	OTU99_13	OTU99_14	OTU99_15	OTU99_16	OTU99_17	OTU99_18	OTU99_19	OTU99_20	OTU99_21	OTU99_22	OTU99_23	OTU99_24	OTU99_25	OTU99_26	OTU99_27	OTU99_28	OTU99_29	OTU99_30	OTU99_31	OTU99_32	OTU99_33	OTU99_34	OTU99_35	OTU99_36	OTU99_37	OTU99_38	OTU99_39	OTU99_40	Number of sequences OTUs 100	
■					■																																			31	OTU100_001
■					■																																			18	OTU100_002
■					■																																			16	OTU100_003
■					■																																			11	OTU100_007
■					■																																			3	OTU100_015
■					■																																			3	OTU100_016
■					■																																			2	OTU100_017
■					■																																			2	OTU100_018
■					■																																			1	OTU100_028
■					■																																			1	OTU100_038
■					■																																			1	OTU100_042
■					■																																			1	OTU100_043
■					■																																			1	OTU100_047
■					■																																			1	OTU100_049
■					■																																			1	OTU100_050
■					■																																			1	OTU100_054
■					■																																			1	OTU100_066
■					■																																			1	OTU100_068
■					■																																			1	OTU100_095
■					■																																			1	OTU100_099
■					■																																			1	OTU100_101
■					■																																			1	OTU100_105
■					■																																			1	OTU100_114
■					■																																			1	OTU100_128
■					■																																			1	OTU100_132
■							■																																		13	OTU100_005
■							■																																		10	OTU100_008
■							■																																		1	OTU100_022
■							■																																		1	OTU100_024
■							■																																		1	OTU100_025
■							■																																		1	OTU100_029
■							■																																		1	OTU100_032
■							■																																		1	OTU100_036
■							■																																		1	OTU100_039


OTUs cut-off 100%
OTUs cut-off 99%
OTUs cut-off 97%
Number of sequences OTUs 100
OTU100_044
OTU100_051
OTU100_057
OTU100_061
OTU100_079
OTU100_086
OTU100_089
OTU100_092
OTU100_096
OTU100_110
OTU100_124
OTU100_065
OTU100_070
OTU100_076
OTU100_112
OTU100_058
OTU100_075
OTU100_040
OTU100_134
OTU100_111
OTU100_048
OTU100_053
OTU100_081
OTU100_135
OTU100_072
OTU100_071
OTU100_060
OTU100_004
OTU100_009
OTU100_012
OTU100_014
OTU100_020
OTU100_021
OTU100_035
OTU100_045
OTU100_056
OTU100_073
OTU100_078
OTU100_080
1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 13 8 4 4 2 1 1 1 1 1 1 1
OTU99_40
OTU99_39
OTU99_38
OTU99_37
OTU99_36
OTU99_35
OTU99_34
OTU99_33
OTU99_32
OTU99_31
OTU99_30
OTU99_29
OTU99_28
OTU99_27
OTU99_26
OTU99_25
OTU99_24
OTU99_23
OTU99_22
OTU99_21
OTU99_20
OTU99_19
OTU99_18
OTU99_17
OTU99_16
OTU99_15
OTU99_14
OTU99_13
OTU99_12
OTU99_11
OTU99_10
OTU99_9
OTU99_8
OTU99_7
OTU99_6
OTU99_5
OTU99_4
OTU99_3
OTU99_2
OTU99_1
OTU97_5
OTU97_4
OTU97_3
OTU97_2
OTU97_1

OTUs cut-off 97%					OTUs cut-off 99%																																				OTUs cut-off 100%	
OTU97_1	OTU97_2	OTU97_3	OTU97_4	OTU97_5	OTU99_1	OTU99_2	OTU99_3	OTU99_4	OTU99_5	OTU99_6	OTU99_7	OTU99_8	OTU99_9	OTU99_10	OTU99_11	OTU99_12	OTU99_13	OTU99_14	OTU99_15	OTU99_16	OTU99_17	OTU99_18	OTU99_19	OTU99_20	OTU99_21	OTU99_22	OTU99_23	OTU99_24	OTU99_25	OTU99_26	OTU99_27	OTU99_28	OTU99_29	OTU99_30	OTU99_31	OTU99_32	OTU99_33	OTU99_34	OTU99_35	OTU99_36	OTU99_37	OTU99_38	OTU99_39	OTU99_40	Number of sequences OTUs 100	
	■					■																																			1	OTU100_088
	■					■																																			1	OTU100_109
	■					■																																			1	OTU100_120
	■					■																																			1	OTU100_121
	■					■																																			1	OTU100_123
	■					■																																			1	OTU100_125
	■							■																																	11	OTU100_006
	■							■																																	5	OTU100_011
	■							■																																	4	OTU100_013
	■							■																																	1	OTU100_027
	■							■																																	1	OTU100_034
	■							■																																	1	OTU100_041
	■							■																																	1	OTU100_052
	■							■																																	1	OTU100_074
	■							■																																	1	OTU100_083
	■							■																																	1	OTU100_084
	■							■																																	1	OTU100_090
	■							■																																	1	OTU100_093
	■							■																																	1	OTU100_097
	■							■																																	1	OTU100_100
	■							■																																	1	OTU100_102
	■							■																																	1	OTU100_107
	■							■																																	1	OTU100_108
	■							■																																	1	OTU100_113
	■							■																																	1	OTU100_115
	■								■																																7	OTU100_010
	■								■																																1	OTU100_106
	■									■																															1	OTU100_059
	■									■																															1	OTU100_077
	■									■																															1	OTU100_122
	■									■																															1	OTU100_126
	■											■																													1	OTU100_023
	■											■																													1	OTU100_026
	■											■																													1	OTU100_033
	■												■																												1	OTU100_055
	■												■																												1	OTU100_064
	■												■																												1	OTU100_067
	■													■																											1	OTU100_082
	■													■																											1	OTU100_118

OTUs cut-off 97%					OTUs cut-off 99%																																								OTUs cut-off 100%	
OTU97_1	OTU97_2	OTU97_3	OTU97_4	OTU97_5	OTU99_1	OTU99_2	OTU99_3	OTU99_4	OTU99_5	OTU99_6	OTU99_7	OTU99_8	OTU99_9	OTU99_10	OTU99_11	OTU99_12	OTU99_13	OTU99_14	OTU99_15	OTU99_16	OTU99_17	OTU99_18	OTU99_19	OTU99_20	OTU99_21	OTU99_22	OTU99_23	OTU99_24	OTU99_25	OTU99_26	OTU99_27	OTU99_28	OTU99_29	OTU99_30	OTU99_31	OTU99_32	OTU99_33	OTU99_34	OTU99_35	OTU99_36	OTU99_37	OTU99_38	OTU99_39	OTU99_40	Number of sequences OTUs 100	
	■													■																															1	OTU100_085
	■														■																														1	OTU100_117
	■														■																														1	OTU100_131
	■																■																												1	OTU100_063
	■																■																												1	OTU100_069
	■																		■																										1	OTU100_094
	■																		■																										1	OTU100_103
	■																			■																									1	OTU100_133
	■																					■																							1	OTU100_030
	■																						■																						1	OTU100_091
	■																								■																				1	OTU100_087
	■																												■																1	OTU100_031
	■																																	■											1	OTU100_116
	■																																		■										1	OTU100_130
	■																																			■									1	OTU100_127
	■																																						■						1	OTU100_129
	■																																								■				1	OTU100_119
	■																																										■		1	OTU100_098
	■																																											■	1	OTU100_104
		■																■																											2	OTU100_019
		■																												■															1	OTU100_046
			■																		■																								1	OTU100_037
				■																													■												1	OTU100_062
161	118	3	1	1	104	43	41	36	8	4	4	3	3	3	2	2	2	2	2	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	284	Number of sequences OTUs 97 and 99

Supplementary references

1. **Maidak BL, Cole JR, Lilburn TG, Parker Jr CT, Saxman PR, Farris RJ, Garrity GM, Olsen GJ, Schmidt TM, Tiedje JM.** 2001. The RDP-II (Ribosomal Database Project). Nucleic Acids Research **29:**173-174.
2. **Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ.** 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Research **25:**3389-3402.
3. **Lane DJ.** 1991 16S/23S rRNA sequencing. In E. Stackebrandt and M. Goodfellow (ed.)., John Willy and Sons, New York.
4. **Weisburg WG, Barns SM, Pelletier DA, Lane DJ.** 1991. 16S ribosomal DNA amplification for phylogenetic study. J Bacteriol **173:**697-703.
5. **Lopez-Siles M, Khan TM, Duncan SH, Harmsen HJ, Garcia-Gil LJ, Flint HJ.** 2012. Cultured representatives of two major phylogroups of human colonic *Faecalibacterium prausnitzii* can utilize pectin, uronic acids, and host-derived substrates for growth. Appl Environ Microbiol **78:**420-428.

Supplemental materials *Article III*

The following supplement accompanies the article

Mucosa-associated *Faecalibacterium prausnitzii* and *Escherichia coli* co-abundance can distinguish Irritable Bowel Syndrome and Inflammatory Bowel Disease phenotypes

Mireia Lopez-Siles, Margarita Martinez-Medina, David Busquets, Miriam Sabat-Mir, Sylvia H. Duncan, Harry J. Flint, Xavier Aldeguer, and L. Jesús Garcia-Gil.

International Journal of Medical Microbiology 2014; 304:464-475.

This supplement includes additional data about the development of the novel qPCR assay described in this work for total *F. prausnitzii* detection and quantification. Additional data of comparisons about groups of patients is also provided.

Supplementary text

Primers and hydrolysis probes design for quantitative PCR (qPCR)

Primers and probes sequences designed in this study are listed in Table 2 on the main text.

Specific primers set and hydrolysis probe targeting the 16S rRNA gene of *F. prausnitzii*

A consensus sequence for *F. prausnitzii* 16S rRNA gene was constructed from using the sequences detailed in Table S1 and further compared against other 16S rRNA sequences of *F. prausnitzii* close relatives. DNA regions exclusive for *F. prausnitzii* were used for primers and hydrolysis probes designs with Primer Express® version 3.0 software (Applied Biosystems, Foster City, CA, USA). Thermodynamic characteristics such as melting temperature, secondary structure and primer-dimer free energy formation as well as G+C content were considered after analysis with NetPrimer® software (available at <http://premierbiosoft.com/netprimer>, PREMIER Biosoft International, California) to choose the best oligonucleotides.

Internal Amplification Control (IAC) design

The IAC consisted of a primers and a hydrolysis probe set, based on a random DNA target sequence, with no identity with any known sequence. Oligonucleotides were designed as described above and the set which showed the shortest amplicon length was chosen.

NetPrimer® analysis software was additionally used to check possible cross dimers interactions between the IAC and *F. prausnitzii*-targeted oligonucleotides.

Reagents optimization of the multiplex quantitative PCR (qPCR)

Experiments were performed using different primer and probe concentrations ranging from 50 to 900 nM to optimize their concentrations in the qPCR assay. To determine the IAC target molecules concentration to include in each assay without competence interference with the qPCR for *F. prausnitzii*, the efficiency of the reaction was analyzed using 10^4 and 10^3 IAC DNA copies. Values were compared with those from reactions run in uniplex fashion. The efficiency of the *F. prausnitzii* qPCR assay was not affected when 10^3 target copies for the IAC were added. Therefore, this IAC's DNA quantity was further used in all qPCR reactions as inhibition control.

Features of the novel qPCR assays for *F. prausnitzii* (two phylogroups)

Specificity tests

The specificity of the oligonucleotides was tested by comparing against the Ribosomal Database Project II (RDP) (Maidak et al., 2001) and GenBank database through Seqmatch and BLAST (Altschul et al., 1997) tools, respectively. The in silico analysis of the oligonucleotide set of choice showed that primer Fpra428F and probe Fpra493PR were unequivocally specific for *F. prausnitzii* and targeted all the isolates available to date, whereas the Fpra583R reverse primer was group-specific. No matching sequences were found for the IAC set, as expected.

For the multiplex *F. prausnitzii* qPCR assay, specificity was also tested in vitro by comparing the quantification of pure *F. prausnitzii* DNA (10 ng) recovered from 9 isolates, representative of both phylogroups. DNAs from 80 additional representative bacterial species (see list on Table S2) which are either close relatives of *F. prausnitzii* or belong to the major groups of bacteria present in the colon were also included. The qPCR reaction was carried out as described in the methods section of the main paper. Negative results were cross checked by no inhibition of the IAC and by alternative amplification by end-point, conventional PCR with universal bacterial primers Bac27F and Uni1492R as previously reported (Lane, 1991 ; Weisburg et al., 1991). Results from the specificity test are also shown in Table S2. The assay was totally specific. All the *F. prausnitzii* isolates were detected and no statistically significant differences in C_q values among them were observed. There was no cross-reaction with any of the non-target microorganisms, and negative results were validated by no delay on the C_q of IAC and positive amplification by conventional PCR.

Sensitivity tests: confident quantification range and detection limit

To determine the confident quantification range of the assay, ten-fold serial dilutions of a linealized plasmid containing a single copy of the 16S rRNA gene of *F. prausnitzii* A2-165 was used. Ten replicas of each dilution were assayed, with a dynamic range from 10^9 to 1 gene copies per reaction. The linear range for quantification was considered for those concentrations having a SD value lower than 0.4 between replicates. The precision of the quantitative experiments was approximately 99%, based on the C_q values for three replicate PCR runs. The standard curve was linear (R^2=0.9986) down to 10^3 16S rDNA copies. The average slope of the linear regression curve over a 7-log range (10^3 to 10^9) was of 3.7, thus the efficiency of the reaction was 85.88%.

To determine the detection limit of the assay, a calibration curve of two-fold serial dilutions between 50 and 10^4 target copies of *F. prausnitzii* 16S rRNA gene was performed. Eight replicas of each dilution were assayed. Data was analyzed by a Probit test (Minitab® 14 Statistical Software, Pennsylvania, USA), in which the ratio of positive/negative amplification events was plotted against the amount of target genes present per reaction. The theoretical minimum number of 16S rRNA genes of *F. prausnitzii* per reaction to have a 95% of probability to obtain a positive detection was 106.6 target genes.

Experiments with 20 mg of tissue specimens spiked with known amounts of *F. prausnitzii* 16S rRNA gene (ranging from 4×10^7 to 0) also showed that mucosal tissue did not inhibit PCR as the fluorescence signal and the C_q value did not differ in the presence or absence of tissue.

Supplementary tables

Table S1. 16S rRNA gene sequences used to perform oligonucleotides design. GenBank accession numbers have been indicated. Sequences from *F. prausnitzii* isolates, related sequences recovered via molecular methods and sequences of the same gene from *F. prausnitzii* close relatives have been included.

Accession number	Characteristics
AJ413954*	*Faecalibacterium prausnitzii* 16S rRNA gene, strain ATCC 27768
X85022*	*F. prausnitzii* DNA for 16S ribosomal RNA, strain ATCC 27766
AJ270470*	Butyrate-producing bacterium L2-6 16S rRNA gene
AJ270469*	Butyrate-producing bacterium A2-165 16S rRNA gene
AY305307*	Butyrate-producing bacterium M21/2 16S ribosomal RNA gene
HQ457025*	*F. prausnitzii* strain S4L/4 16S ribosomal RNA gene
HQ457024*	*F. prausnitzii* strain S3L/3 16S ribosomal RNA gene
JN037415*	*F. prausnitzii* strain L2-15 16S ribosomal RNA gene
JN037416*	*F. prausnitzii* strain L2-39 16S ribosomal RNA gene
JN037417*	*F. prausnitzii* strain L2-61 16S ribosomal RNA gene
AM075671*	Uncultured bacterium partial 16S rRNA gene, isolate C0703§.
AM075691*	Uncultured bacterium partial 16S rRNA gene, isolate C0402§.
AM075696*	Uncultured bacterium partial 16S rRNA gene, isolate C0801§.
AM075683*	Uncultured bacterium partial 16S rRNA gene, isolate C1403§.
AM075730*	Uncultured bacterium partial 16S rRNA gene, isolate CD1902§.
AM075738*	Uncultured bacterium partial 16S rRNA gene, isolate UC0102§.
AY169429*	*Faecalibacterium prausnitzii* clone 1-84 16S ribosomal RNA gene, partial sequence
AY169430*	*Faecalibacterium prausnitzii* clone 1-88 16S ribosomal RNA gene, partial sequence
AY169427*	*Faecalibacterium prausnitzii* clone 1-79 16S ribosomal RNA gene, partial sequence
AF132237*	Uncultured bacterium adhufec13 16S ribosomal RNA gene, partial sequence§
AF132236*	Uncultured bacterium adhufec113 16S ribosomal RNA gene, partial sequence§
AF132246*	Uncultured bacterium adhufec218 16S ribosomal RNA gene, partial sequence§
AF132265*	Uncultured bacterium adhufec365 16S ribosomal RNA gene, partial sequence§
AF153871*	Uncultured bacterium adhufec08.25 16S ribosomal RNA gene, partial sequence§
AY494671*	Uncultured *Faecalibacterium* sp. clone FIRM8 16S ribosomal RNA gene, partial sequence
X98011	*Anaerofilum agile* 16S rRNA gene
X97852	*Anaerofilum pentosovorans* 16S rRNA gene
L09177	*Clostridium cellulosi* 16S ribosomal RNA (16S rRNA) gene
M59095	*Clostridium leptum* 16S ribosomal RNA
AJ305238	*Clostridium leptum*; DSM 753T
M59116	*Clostridium sporosphaeroides* 16S ribosomal RNA
X66002	*Clostridium sporosphaeroides*; DSM 1294
X81125	*Clostridium viride* 16S rRNA gene
L34618	*Eubacterium desmolans* 16S ribosomal RNA
L34625	*Eubacterium siraeum* 16S ribosomal RNA
AY445600	*Ruminococcus albus* strain 7 16S ribosomal RNA gene, complete
AY445594	*Ruminococcus albus* strain 8 16S ribosomal RNA gene, complete
AY445592	*Ruminococcus albus* strain B199 16S ribosomal RNA gene, complete
AY445596	*Ruminococcus albus* strain KF1 16S ribosomal RNA gene, complete
AY445602	*Ruminococcus albus* strain RO13 16S ribosomal RNA gene, complete
X85099	*Ruminococcus bromii* 16S rRNA gene
L76600	*Ruminococcus bromii* small subunit ribosomal RNA (16S rDNA) gene
X85100	*Ruminococcus callidus* 16S rRNA gene
L76596	*Ruminococcus callidus* small subunit ribosomal RNA (16S rDNA)
AM915269	*Ruminococcus flavefaciens* partial 16S rRNA gene, type strain C94T=ATCC19208
AF030449	*Ruminococcus flavefaciens* strain ATCC 49949 16S ribosomal RNA, partial sequence
AY445599	*Ruminococcus flavefaciens* strain B146 16S ribosomal RNA gene, complete sequence
AY445597	*Ruminococcus flavefaciens* strain FD1 16S ribosomal RNA gene, complete sequence
AY445595	*Ruminococcus flavefaciens* strain JM1 16S ribosomal RNA gene, complete sequence
AY445603	*Ruminococcus flavefaciens* strain LB4 16S ribosomal RNA gene, complete sequence
AY445598	*Ruminococcus flavefaciens* strain R13e2 16S ribosomal RNA gene, complete sequence

* Sequences used to obtain the *F. prausnitzii* 16S rRNA gene consensus sequence for oligonucleotides design,
§ Sequence of the genus *Faecalibacterium*

Table S2. Growth conditions and source of the bacterial strains used in this study. The results obtained from the specificity tests are also included.

Source of DNA information		Growth (2)		Specificity test information			
Phylogeny	Strain/source (1)	Media	T(ºC)	ng (3)	cnPCR	IAC	qPCR
Firmicutes							
Faecalibacterium prausnitzii ATCC 27768^T	ATCC 27768	M2GSC	37	10	+	+	+
F. prausnitzii A2-165	DSM17677	M2GSC	37	10	+	+	+
F. prausnitzii M21/2	nd	M2GSC	37	10	+	+	+
F. prausnitzii L2-15	nd	M2GSC	37	10	+	+	+
F. prausnitzii L2-39	nd	M2GSC	37	10	+	+	+
F. prausnitzii L2-6	nd	M2GSC	37	10	+	+	+
F. prausnitzii L2-61	nd	M2GSC	37	10	+	+	+
F. prausnitzii S3L/3	nd	M2GSC	37	10	+	+	+
F. prausnitzii S4L/4	nd	M2GSC	37	10	+	+	+
Anaerofilum agile	DSM4272	nc	nc	1.6	+	+	-
Eubacterium siraeum	DSMZ 15702	nc	nc	6.9	+	+	-
Eubacterium halii	DSMZ 17630	nc	nc	1	+	+	-
Clostridium viride	DSM6836	nc	nc	10	+	+	-
Clostridium leptum	DSM753	nc	nc	10	+	+	-
Ruminococcus albus	DSM20455	nc	nc	10	+	+	-
Clostridium acetobutylicum	CECT 979	AN	37	3.7	+	+	-
Clostridium botulinum type E	CECT4611	LiB	37	10	+	+	-
Bacillus cereus	NCTC11145	AN	30	10	+	+	-
Bacillus megaterium	DSM319	AN	30	10	+	+	-
Bacillus sp.	CECT 40	AN	30	10	+	+	-
Bacillus subtilis	NCTC10400	AN	30	2.3	+	+	-
Bacillus subtilis sups. spizizwnii	CECT 482	AN	30	10	+	+	-
Listeria grayi	CECT931	BHI	37	10	+	+	-
Listeria innocua	CECT 910	BHI	37	10	+	+	-
Paenibacillus polymyxa	DSM372	BHI	37	2.1	+	+	-
Staphylococcus aureus	ATCC9144	AN	37	10	+	+	-
Staphylococcus epidermidis	CECT 231	AN	37	10	+	+	-
Enterococcus avium	CECT 968	BHI	37	10	+	+	-
Enterococcus columbae	CECT 4798	BHI	37	10	+	+	-
Enterococcus durans	CECT 411	BHI	37	10	+	+	-
Enterococcus faecalis	CECT 481	BHI	37	10	+	+	-
Enterococcus faecium	CECT 410	BHI	37	10	+	+	-
Enterococcus gallinarum	CECT 970	BHI	37	10	+	+	-
Enterococcus mundtii	CECT 972	BHI	37	10	+	+	-
Lactobacillus acidophilus	CECT 903	MRS	30	6.3	+	+	-
Lactococcus lactis	CECT 185	MRS	30	3.8	+	+	-
Streptococcus agalactiae	CECT 183	BHI	37	7.2	+	+	-
Streptococcus anginosus	CECT 948	BHI	37	10	+	+	-
Streptococcus equi subsp. *equi*	CECT 989	BHI	37	10	+	+	-
Streptococcus equinus	CECT 213	BHI	37	10	+	+	-
Streptococcus intermedius	CECT 803	BHI	37	10	+	+	-
Streptococcus mutans	CECT 479	BHI	37	3.8	+	+	-
Streptococcus oralis	CECT 907	BHI	37	10	+	+	-
Streptococcus pneumoniae	CECT 993	BHI	37	10	+	+	-
Streptococcus pyogenes	CECT 598	BHI	37	10	+	+	-
Streptococcus salivarus	CECT 805	BHI	37	10	+	+	-
Streptococcus sanguinis	CECT 480	BHI	37	5.5	+	+	-
Streptococcus sobrinus	CECT 4034	BHI	37	6.5	+	+	-
Streptococcus suis	CECT 958	BHI	37	10	+	+	-
Streptococcus thermophilus	CECT 986	BHI	37	10	+	+	-
Streptococcus uberis	CECT 994	BHI	37	10	+	+	-

Source of DNA information		Growth (2)		Specificity test information			
Phylogeny	Strain/source [1]	Media	T(ºC)	ng [3]	cnPCR	IAC	qPCR
Actinobacteria							
Corynebacterium bovis	DSM20582	MRS	37	4.8	+	+	-
Kocuria rhizophila	DSM348	AN	30	2.3	+	+	-
Micrococcus luteus	CECT 241	AN	30	2.6	+	+	-
Mycobacterium phlei	CECT 3009	BHI	37	10	+	+	-
Streptomyces griseus	DSM40236	PDA	30	10	+	+	-
Bifidobacterium adolescentis	CECT 5781	AN	37	0.4	+	+	-
Bifidobacterium breve	CECT 4839	AN	37	2.0	+	+	-
Bacteroidetes							
Bacteroides fragilis	DSMZ 2151	nc	nc	10	+	+	-
Bacteroides uniformis	DSMZ 6597	nc	nc	10	+	+	-
Bacteroides vulgatus	DSMZ 1447	nc	nc	10	+	+	-
Proteobacteria							
Methylophilus methylotrophus	DSM5691	CZ	30	10	+	+	-
Campylobacter jejuni	DSM4688	BA	37	10	+	+	-
Citrobacter freundii	CECT 401	AN	30	10	+	+	-
Enterobacter aerogenes	CECT 684	AN	30	10	+	+	-
Enterobacter cloacae	CECT 194	AN	30	10	+	+	-
Enterobacter sakazakii	CECT 858	AN	30	10	+	+	-
Enterobacter sakazakii	ATCC51329	AN	30	0.4	+	+	-
Enterobacter amnigenus (Sakazakii)	CECT 4078	AN	37	10	+	+	-
Enterobacter gergoviae (Sakazakii)	CECT 857	AN	37	10	+	+	-
Escherichia coli	CECT 100	AN	37	10	+	+	-
Escherichia coli	CECT 101	AN	37	10	+	+	-
Escherichia coli	CECT 105	AN	37	10	+	+	-
Escherichia coli	CECT 12242	AN	37	10	+	+	-
Escherichia coli	CECT 831	AN	37	10	+	+	-
Escherichia coli	CECT 4201	AN	37	10	+	+	-
Escherichia coli	CECT 4084	AN	37	10	+	+	-
Escherichia coli	CECT 405	AN	37	10	+	+	-
Escherichia coli	ATCC10536	AN	37	10	+	+	-
Klebsiella pneumoniae ssp. pneumoniae	CECT 143	AN	37	10	+	+	-
Proteus mirabilis	CECT 170	AN	37	10	+	+	-
Salmonella LT2	CECT878	AN	37	10	+	+	-
Salmonella TA98	CECT880	AN	37	10	+	+	-
Serratia marcescens	CECT846	AN	25	10	+	+	-
Shigella sonnei	CECT457	AN	37	10	+	+	-
Pseudomonas aeruginosa	CECT 532	AN	30	10	+	+	-
Pseudomonas fluorescens	CECT 378	AN	30	10	+	+	-
Pseudomonas mendocina	CECT320	AN	30	10	+	+	-
Pseudomonas putida	CECT 324	AN	30	4.1	+	+	-
Human	Eurogentec						

[1] ATCC: American Type Culture Collection (Manassas, VA, USA); CECT: Colección Española de Cultivos Tipo (Valencia, Spain); DSMZ: Deutche Sammlung von Mikroorganismen and Zellkulturen (Braunschweig, Germany), NCTC: National Collection of Type Cultures (London,UK), nd: not deposited (stocks held by the authors, Rowett Institute of nutrition and Health, Aberdeen, United Kingdom).

[2] nc: not cultured. BHI (Brain Heart Infusion Broth), AN (Nutrient Agar), BA (Blood Agar), MRS (Man, Rogosa and Sharpe medium), LiB(Liver Broth, CECT medium #15), CZ (Colby and Zathman medium, DSMZ medium #606), PDA (Potato Dextrose Agar), M2GSC (modified Med2 of Hobson, (Lopez-Siles et al., 2012))

[3] ng of genomic DNA used for the specificity test. When possible, 10ng was used. The DNA was obtained from 1ml of bacterial culture at the stationary growth phase or for nc strains, the dried culture directly obtained from the culture type collection was rehydrated with the appropriate buffer for DNA extraction and used to DNA purification.

Table S3. *F. prausnitzii* and *E. coli* abundances (median ± SD) categorized by activity status (A, active; I, inactive) in Inflammatory Bowel Disease patients by disease phenotype (I-CD, ileal CD; IC-CD, ileocolonic CD, C-CD, colonic CD).

	Activity	N patients *(N biopsies)*	*F. prausnitzii*		*E. coli*	
Proctitis	A	5 *(15)*	5.07±0.35	n.s	3.46±0.85	n.s
	I	1 *(3)*	5.14±0.06		3.03±0.29	
Distal UC	A	9 *(27)*	4.44±0.68	n.s	2.92±1.13	n.s
	I	2 *(6)*	4.42±0.12		2.13±1.00	
Extensive UC	A	6 *(14)*	5.34±0.67	n.s	4.62±1.25	n.s
	I	1 *(1)*	4.32		0.94	
C-CD	A	5 *(9)*	5.10±0.72	n.s	4.44±0.93	n.s
	I	9 *(19)*	5.06±1.01		4.61±0.92	
IC-CD	A	7 *(11)*	4.44±1.22	n.s	5.04±1.31	p=0.006
	I	8 *(14)*	4.41±0.86		3.56±0.85	
I-CD	A	8 *(11)*	3.96±0.93	n.s	4.96±1.12	p=0.050
	I	14 *(29)*	3.52±1.51		4.50±1.09	

Supplementary references

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., Lipman, D.J., 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Research 25, 3389-3402.

Lane, D.J., 1991 16S/23S rRNA sequencing. In E. Stackebrandt and M. Goodfellow (ed.)., John Willy and Sons, New York.

Lopez-Siles, M., Khan, T.M., Duncan, S.H., Harmsen, H.J., Garcia-Gil, L.J., Flint, H.J., 2012. Cultured representatives of two major phylogroups of human colonic *Faecalibacterium prausnitzii* can utilize pectin, uronic acids, and host-derived substrates for growth. Appl Environ Microbiol 78, 420-428.

Maidak, B.L., Cole, J.R., Lilburn, T.G., Parker Jr, C.T., Saxman, P.R., Farris, R.J., Garrity, G.M., Olsen, G.J., Schmidt, T.M., Tiedje, J.M., 2001. The RDP-II (Ribosomal Database Project). Nucleic Acids Research 29, 173-174.

Weisburg, W.G., Barns, S.M., Pelletier, D.A., Lane, D.J., 1991. 16S ribosomal DNA amplification for phylogenetic study. J Bacteriol 173, 697-703.

Supplemental materials *Article IV*

The following supplement accompanies the article

Changes in the abundance of *Faecalibacterium prausnitzii* phylogroups I and II in the intestinal mucosa of inflammatory bowel disease and patients with colorectal cancer

Mireia Lopez-Siles, Margarita Martinez-Medina, Romà Surís-Valls, Xavier Aldeguer, Miriam Sabat-Mir, Sylvia H. Duncan, Harry J. Flint, and L. Jesús Garcia-Gil.

Accepted in Inflammatory Bowel Disease on 3rd August 2015, doi:10.1097/MIB.0000000000000590

This supplement includes additional data about the development of the novel qPCR assay described in this work for the simultaneous and differential detection and quantification of *F. prausnitzii* phylogroups. Additional data of comparisons about groups of patients is also provided.

Supplementary text

Bacterial strains, growth conditions and DNA extraction from pure cultures

F. prausnitzii strains were from stocks held by the authors (Rowett Institute of Nutrition and Health, Aberdeen, United Kingdom) and several came from previous studies (1-5). Additional bacterial strains were either available in our laboratory collection or were otherwise obtained from several biological resource centres specified in Table S2. When possible, bacteria were cultured aerobically or anaerobically on the recommended medium. DNA was extracted and purified by using the Wizard™ Genomic Purification Kit (Promega Corporation, USA) following the manufacturer's guidelines.

Optimisation and characterization of the multiplex qPCR assay for *F. prausnitzii* phylogroups

To determine the best reagent concentrations for the qPCR assay, experiments were performed using different primer and probe concentrations ranging from 50 to 900 nM. Those reagents concentrations that yield the maximum fluorescent signal and the lowest quantification cycle (C_q) value for 10^6 copies/reaction of the target DNA were chosen as optimal, and have therefore been used for further quantification in samples (as described in the main text section Quantitative PCR conditions).

Inclusivity and exclusivity tests

For the multiplex quantification of *F. prausnitzii* phylogroups using a qPCR assay, specificity was also tested *in vitro* by comparing the quantification of pure *F. prausnitzii* DNA (10 ng) recovered from nine isolates, representative of both phylogroups. DNA from 80 additional representative bacterial species (see list in Table S2) which are either close relatives of *F. prausnitzii* or belong to the major groups of bacteria present in the colon were also included.

The qPCR was carried out as described in the section Quantitative PCR conditions of the main text. Negative results were cross checked by alternative amplification by end-point, conventional PCR with universal bacterial primers Bac27F and Uni1492R as previously reported (6, 7). Results from the specificity test are also shown in Table S2. The assay was totally specific. All the *F. prausnitzii* isolates were only detected for the phylogroup they belong to, and no statistically significant differences in C_q values between isolates were observed. There was no cross-reaction with any of the non-target microorganisms, and negative results were validated by a positive amplification by conventional PCR.

Linear quantification range and efficiency of the qPCR

To determine the confident quantification range of the assay, decaplicate ten-fold dilutions (ranging from 2×10^8 to 2 target gene copies per reaction) of a linearized plasmid containing either a single copy of the 16S rRNA gene of *F. prausnitzii* S3L/3 (phylogroup I) or *F. prausnitzii* DSM 17677 (phylogroup II) were used. The linear range for quantification was considered for those concentrations having a SD value lower than 0.34 between replicates. Regression analysis plotting the obtained Cq against the logarithm of the number of target genes in the reaction was also performed. The efficiency of the qPCR assay was calculated using the formula: Efficiency= $[10^{(-1/\text{slope})}]$-1.

Detection limit of the assay

A calibration curve of two-fold serial dilutions between 1 and 100 target copies of *F. prausnitzii* 16S rRNA gene was performed. Eight replicas of each dilution were assayed. Data was analyzed by a Probit test (Minitab® 14 Statistical Software, Pennsylvania, USA), in which the ratio of positive/negative amplification events was plotted against the amount of target genes present per reaction.

Supplementary tables

Table S1. 16S rRNA gene sequences used to perform oligonucleotide design. GenBank accession numbers have been indicated. Sequences from *F. prausnitzii* isolates, related sequences recovered via molecular methods and sequences of the same gene from *F. prausnitzii* close relatives have been included.

Accession number	Characteristics
AJ413954*[1]	*Faecalibacterium prausnitzii* 16S rRNA gene, strain ATCC 27768
X85022*[1]	*F. prausnitzii* DNA for 16S ribosomal RNA, strain ATCC 27766
AY305307*[1]	Butyrate-producing bacterium M21/2 16S ribosomal RNA gene
HQ457025*[1]	*F. prausnitzii* strain S4L/4 16S ribosomal RNA gene
HQ457024*[1]	*F. prausnitzii* strain S3L/3 16S ribosomal RNA gene
AJ270469*[2]	Butyrate-producing bacterium A2-165 16S rRNA gene
AJ270470*[2]	Butyrate-producing bacterium L2-6 16S rRNA gene
JN037415*[2]	*F. prausnitzii* strain L2-15 16S ribosomal RNA gene
JN037416*[2]	*F. prausnitzii* strain L2-39 16S ribosomal RNA gene
JN037417*[2]	*F. prausnitzii* strain L2-61 16S ribosomal RNA gene
HQ457026*[2]	*F. prausnitzii* strain HTF-A 16S ribosomal RNA gene
HQ457027*[2]	*F. prausnitzii* strain HTF-B 16S ribosomal RNA gene
HQ457028*[2]	*F. prausnitzii* strain HTF-C 16S ribosomal RNA gene
HQ457029*[2]	*F. prausnitzii* strain HTF-E 16S ribosomal RNA gene
HQ457030*[2]	*F. prausnitzii* strain HTF-F 16S ribosomal RNA gene
HQ457031*[2]	*F. prausnitzii* strain HTF-I 16S ribosomal RNA gene
HQ457032*[2]	*F. prausnitzii* strain HTF-60C 16S ribosomal RNA gene
HQ457033*[2]	*F. prausnitzii* strain HTF-75H 16S ribosomal RNA gene
AY169429*	*Faecalibacterium prausnitzii* clone 1-84 16S ribosomal RNA gene, partial sequence
AY169430*	*Faecalibacterium prausnitzii* clone 1-88 16S ribosomal RNA gene, partial sequence
AY169427*	*Faecalibacterium prausnitzii* clone 1-79 16S ribosomal RNA gene, partial sequence
AF132237*	Uncultured bacterium adhufec13 16S ribosomal RNA gene, partial sequence[§]
AF132236*	Uncultured bacterium adhufec113 16S ribosomal RNA gene, partial sequence[§]
AF132246*	Uncultured bacterium adhufec218 16S ribosomal RNA gene, partial sequence[§]
AF132265*	Uncultured bacterium adhufec365 16S ribosomal RNA gene, partial sequence[§]
AY494671*	Uncultured *Faecalibacterium* sp. clone FIRM8 16S ribosomal RNA gene, partial sequence
EF205929*	Uncultured bacterium clone 46706[§]
EF205662*	Uncultured bacterium clone 58014[§]
EF206222*	Uncultured bacterium clone 56806[§]
EF206249*	Uncultured bacterium clone 57601[§]
EF205881*	Uncultured bacterium clone 35509[§]
EF205761*	Uncultured bacterium clone 59415[§]
EF205681*	Uncultured bacterium clone 58033[§]
X98011	*Anaerofilum agile* 16S rRNA gene
X97852	*Anaerofilum pentosovorans* 16S rRNA gene
L09177	*Clostridium cellulosi* 16S ribosomal RNA (16S rRNA) gene
M59095	*Clostridium leptum* 16S ribosomal RNA
AJ305238	*Clostridium leptum*; DSM 753T
M59116	*Clostridium sporosphaeroides* 16S ribosomal RNA
X66002	*Clostridium sporosphaeroides*; DSM 1294
X81125	*Clostridium viride* 16S rRNA gene
L34618	*Eubacterium desmolans* 16S ribosomal RNA
L34625	*Eubacterium siraeum* 16S ribosomal RNA
AY445600	*Ruminococcus albus* strain 7 16S ribosomal RNA gene, complete
AY445594	*Ruminococcus albus* strain 8 16S ribosomal RNA gene, complete
AY445592	*Ruminococcus albus* strain B199 16S ribosomal RNA gene, complete
AY445596	*Ruminococcus albus* strain KF1 16S ribosomal RNA gene, complete
AY445602	*Ruminococcus albus* strain R013 16S ribosomal RNA gene, complete
X85099	*Ruminococcus bromii* 16S rRNA gene
L76600	*Ruminococcus bromii* small subunit ribosomal RNA (16S rDNA) gene
X85100	*Ruminococcus callidus* 16S rRNA gene

Accession number	Characteristics
L76596	*Ruminococcus callidus* small subunit ribosomal RNA (16S rDNA)
AM915269	*Ruminococcus flavefaciens* partial 16S rRNA gene, type strain C94T=ATCC19208
AF030449	*Ruminococcus flavefaciens* strain ATCC 49949 16S ribosomal RNA, partial sequence
AY445599	*Ruminococcus flavefaciens* strain B146 16S ribosomal RNA gene, complete sequence
AY445597	*Ruminococcus flavefaciens* strain FD1 16S ribosomal RNA gene, complete sequence
AY445595	*Ruminococcus flavefaciens* strain JM1 16S ribosomal RNA gene, complete sequence
AY445593	*Ruminococcus flavefaciens* strain C94 16S ribosomal RNA gene, complete sequence
AY445603	*Ruminococcus flavefaciens* strain LB4 16S ribosomal RNA gene, complete sequence
AY445601	*Ruminococcus flavefaciens* strain JF1 16S ribosomal RNA gene, complete sequence
AY445598	*Ruminococcus flavefaciens* strain R13e2 16S ribosomal RNA gene, complete sequence

* Sequences used to obtain the *F. prausnitzii* 16S rRNA gene consensus sequence for oligonucleotides design
[1] Sequences used to obtain the *F. prausnitzii* phylogroup I 16S rRNA gene consensus sequence for specific hydrolysis probe design
[2] Sequences used to obtain the *F. prausnitzii* phylogroup II 16S rRNA gene consensus sequence for specific hydrolysis probe design
§ Sequence of the genus *Faecalibacterium*

Table S2. Growth conditions and source of the bacterial strains used in this study. The results obtained from the specificity tests are also included.

Source of DNA information*		Growth (2)		Specificity test information			
Phylogeny	Strain/source (1)	Media	T(ºC)	ng (3)	cnPCR	qPHG1	qPHG2
Firmicutes							
Faecalibacterium prausnitzii ATCC 27768^T	ATCC 27768	M2GSC	37	10	+	+	-
F. prausnitzii M21/2	nd	M2GSC	37	10	+	+	-
F. prausnitzii S3L/3	nd	M2GSC	37	10	+	+	-
F. prausnitzii S4L/4	nd	M2GSC	37	10	+	+	-
F. prausnitzii A2-165	DSM17677	M2GSC	37	10	+	-	+
F. prausnitzii L2-15	nd	M2GSC	37	10	+	-	+
F. prausnitzii L2-39	nd	M2GSC	37	10	+	-	+
F. prausnitzii L2-6	nd	M2GSC	37	10	+	-	+
F. prausnitzii L2-61	nd	M2GSC	37	10	+	-	+
Anaerofilum agile	DSM4272	nc	nc	1.6	+	-	-
Eubacterium siraeum	DSM15702	nc	nc	6.9	+	-	-
Eubacterium halii	DSM17630	nc	nc	1	+	-	-
Clostridium viride	DSM6836	nc	nc	10	+	-	-
Clostridium leptum	DSM753	nc	nc	10	+	-	-
Ruminococcus albus	DSM20455	nc	nc	10	+	-	-
Clostridium acetobutylicum	CECT 979	AN	37	3.7	+	-	-
Clostridium botulinum type E	CECT4611	LiB	37	10	+	-	-
Bacillus cereus	NCTC11145	AN	30	10	+	-	-
Bacillus megaterium	DSM319	AN	30	10	+	-	-
Bacillus sp.	CECT 40	AN	30	10	+	-	-
Bacillus subtilis	NCTC10400	AN	30	2.3	+	-	-
Bacillus subtilis sups. spizizwnii	CECT 482	AN	30	10	+	-	-
Listeria grayi	CECT931	BHI	37	10	+	-	-
Listeria innocua	CECT 910	BHI	37	10	+	-	-
Paenibacillus polymyxa	DSM372	BHI	37	2.1	+	-	-
Staphylococcus aureus	ATCC9144	AN	37	10	+	-	-
Staphylococcus epidermidis	CECT 231	AN	37	10	+	-	-
Enterococcus avium	CECT 968	BHI	37	10	+	-	-
Enterococcus columbae	CECT 4798	BHI	37	10	+	-	-
Enterococcus durans	CECT 411	BHI	37	10	+	-	-
Enterococcus faecalis	CECT 481	BHI	37	10	+	-	-
Enterococcus faecium	CECT 410	BHI	37	10	+	-	-
Enterococcus gallinarum	CECT 970	BHI	37	10	+	-	-
Enterococcus mundtii	CECT 972	BHI	37	10	+	-	-
Lactobacillus acidophilus	CECT 903	MRS	30	6.3	+	-	-
Lactococcus lactis	CECT 185	MRS	30	3.8	+	-	-
Streptococcus agalactiae	CECT 183	BHI	37	7.2	+	-	-
Streptococcus anginosus	CECT 948	BHI	37	10	+	-	-
Streptococcus equi subsp. *equi*	CECT 989	BHI	37	10	+	-	-
Streptococcus equinus	CECT 213	BHI	37	10	+	-	-
Streptococcus intermedius	CECT 803	BHI	37	10	+	-	-
Streptococcus mutans	CECT 479	BHI	37	3.8	+	-	-
Streptococcus oralis	CECT 907	BHI	37	10	+	-	-
Streptococcus pneumoniae	CECT 993	BHI	37	10	+	-	-
Streptococcus pyogenes	CECT 598	BHI	37	10	+	-	-
Streptococcus salivarus	CECT 805	BHI	37	10	+	-	-
Streptococcus sanguinis	CECT 480	BHI	37	5.5	+	-	-
Streptococcus sobrinus	CECT 4034	BHI	37	6.5	+	-	-
Streptococcus suis	CECT 958	BHI	37	10	+	-	-
Streptococcus thermophilus	CECT 986	BHI	37	10	+	-	-
Streptococcus uberis	CECT 994	BHI	37	10	+	-	-

Source of DNA information*		Growth [2]		Specificity test information			
Phylogeny	Strain/source [1]	Media	T(ºC)	ng [3]	cnPCR	qPHG1	qPHG2
Actinobacteria							
Corynebacterium bovis	DSM20582	MRS	37	4.8	+	-	-
Kocuria rhizophila	DSM348	AN	30	2.3	+	-	-
Micrococcus luteus	CECT 241	AN	30	2.6	+	-	-
Mycobacterium phlei	CECT 3009	BHI	37	10	+	-	-
Streptomyces griseus	DSM40236	PDA	30	10	+	-	-
Bifidobacterium adolescentis	CECT 5781	AN	37	0.4	+	-	-
Bifidobacterium breve	CECT 4839	AN	37	2.0	+	-	-
Bacteroidetes							
Bacteroides fragilis	DSM2151	nc	nc	10	+	-	-
Bacteroides uniformis	DSM6597	nc	nc	10	+	-	-
Bacteroides vulgatus	DSM1447	nc	nc	10	+	-	-
Proteobacteria							
Methylophilus methylotrophus	DSM5691	CZ	30	10	+	-	-
Campylobacter jejuni	DSM4688	BA	37	10	+	-	-
Citrobacter freundii	CECT 401	AN	30	10	+	-	-
Enterobacter aerogenes	CECT 684	AN	30	10	+	-	-
Enterobacter cloacae	CECT 194	AN	30	10	+	-	-
Enterobacter sakazakii	CECT 858	AN	30	10	+	-	-
Enterobacter sakazakii	ATCC51329	AN	30	0.4	+	-	-
Enterobacter amnigenus (Sakazakii)	CECT 4078	AN	37	10	+	-	-
Enterobacter gergoviae (Sakazakii)	CECT 857	AN	37	10	+	-	-
Escherichia coli	CECT 100	AN	37	10	+	-	-
Escherichia coli	CECT 101	AN	37	10	+	-	-
Escherichia coli	CECT 105	AN	37	10	+	-	-
Escherichia coli	CECT 12242	AN	37	10	+	-	-
Escherichia coli	CECT 831	AN	37	10	+	-	-
Escherichia coli	CECT 4201	AN	37	10	+	-	-
Escherichia coli	CECT 4084	AN	37	10	+	-	-
Escherichia coli	CECT 405	AN	37	10	+	-	-
Escherichia coli	ATCC10536	AN	37	10	+	-	-
Klebsiella pneumoniae ssp. pneumoniae	CECT 143	AN	37	10	+	-	-
Proteus mirabilis	CECT 170	AN	37	10	+	-	-
Salmonella LT2	CECT878	AN	37	10	+	-	-
Salmonella TA98	CECT880	AN	37	10	+	-	-
Serratia marcescens	CECT846	AN	25	10	+	-	-
Shigella sonnei	CECT457	AN	37	10	+	-	-
Pseudomonas aeruginosa	CECT 532	AN	30	10	+	-	-
Pseudomonas fluorescens	CECT 378	AN	30	10	+	-	-
Pseudomonas mendocina	CECT320	AN	30	10	+	-	-
Pseudomonas putida	CECT 324	AN	30	4.1	+	-	-

* Specificity test with human Xsomal DNA (Eurogentec, Belgium) was also performed

[1] ATCC: American Type Culture Collection (Manassas, VA, USA); CECT: Colección Española de Cultivos Tipo (Valencia, Spain); DSMZ: Deutche Sammlung von Mikroorganismen and Zellkulturen (Braunschweig, Germany), NCTC: National Collection of Type Cultures (London,UK), nd: not deposited (stocks held by the authors, Rowett Institute of nutrition and Health, Aberdeen, United Kingdom).

[2] nc: not cultured. BHI (Brain Heart Infusion Broth), AN (Nutrient Agar), BA (Blood Agar), MRS (Man, Rogosa and Sharpe medium), LiB(Liver Broth, CECT medium #15), CZ (Colby and Zathman medium, DSMZ medium #606), PDA (Potato Dextrose Agar), M2GSC (modified Med2 of Hobson, (1))

[3] ng of genomic DNA used for the inclusivity/exclusivity test. When possible, 10ng was used. The DNA was obtained from 1ml of bacterial culture at the stationary growth phase or for nc strains, the dried culture directly obtained from the culture type collection was rehydrated with the appropriate buffer for DNA extraction and used for DNA purification.

Table S3. Quantification of *F. prausnitzii* phylogroups quantifications when DNA from both were present in the same sample.

Phylogroup I: Phylogroup II proportion	Phylogroup I		Phylogroup II	
	Expected quantity*	Measured quantity*±SD	Expected quantity*	Measured quantity*±SD
100:0		$1.15\times10^{9}\pm1.46\times10^{8}$	0	nd
75:25	7.50×10^{8}	$7.85\times10^{8}\pm1.51\times10^{8}$	1.05×10^{8}	$9.70\times10^{7}\pm1.29\times10^{7}$
50:50	5.00×10^{8}	$4.82\times10^{8}\pm8.45\times10^{7}$	2.09×10^{8}	$2.04\times10^{8}\pm1.56\times10^{7}$
25:75	2.00×10^{8}	$1.80\times10^{8}\pm2.56\times10^{7}$	3.14×10^{8}	$3.07\times10^{8}\pm4.39\times10^{7}$
0:100	0	nd		$4.18\times10^{8}\pm4.73\times10^{7}$

* Mean 16S rRNA gene copies of *F. prausnitzii*
nd, not detected; SD, standard deviation

Table S4. Spearman correlation between mucosa-associated *F. prausnitzii* phylogroups with respect to *E. coli* abundances (16S rRNA gene copies/million bacterial 16S rRNA gene copies) in control (H), Irritable Bowel Syndrome (IBS), Ulcerative Colitis (UC), Crohn's disease (CD) and colorectal cancer (CRC) patients. Inflammatory bowel disease patients have also been depicted by disease location.

Patients	N patients (*N biopsies*)	*Phylogroup I* vs. *E. coli*		*Phylogroup II* vs. *E. coli*	
		ρ	P	ρ	P
H	31 *(48)*	-0.031	0.837	0.109	0.466
IBS	9 *(19)*	0.244	0.314	-0.110	0.966
UC	25 *(50)*	0.041	0.778	-0.061	0.673
CD	45 *(63)*	-0.022	0.861	-0.222	0.081
CRC	20 *(20)*	0.091	0.703	-0.086	0.717
Ulcerative proctitis (E1)	6 *(14)*	-0.277	0.180	-0.366	0.072
Distal UC (E2)	11 *(22)*	-0.351	0.109	-0.381	0.080
Extensive UC or ulcerative pancolitis (E3)	6 *(10)*	0.614	0.059	-0.274	0.444
Ileal-CD (L1)	19 *(25)*	-0.277	0.180	-0.366	0.072
Colonic-CD (L2)	11 *(17)*	0.231	0.373	-0.027	0.918
Ileocolonic-CD (L3)	14 *(18)*	0.133	0.598	-0.240	0.336

Table S5. *F. prausnitzii* and its phylogroups abundance in inflammatory bowel disease patients by disease activity status. Active CD and UC were defined by a CDAI of >150 (47) and a Mayo score >3, respectively.

Diagnostics§	N	*F. prausnitzii**	p-value	Phylogroup I*	p-value	Phylogroup II*	p-value
UC							
active	41	4.80±0.41	0.344	2.62±1.32	0.720	2.92±1.02	0.623
inactive	8	5.02±0.66		2.69±0.78		3.18±0.87	
CD							
active	41	4.31±1.10	0.507	0.61±1.51	0.106	1.50±1.63	0.624
inactive	22	4.25±1.46		1.36±1.80		1.69±1.14	

* Median $\log_{10}$ 16S rRNA gene copies/ million bacterial 16S rRNA gene copies ± standard deviations
§UC, ulcerative colitis; CD, Crohn's disease

Table S6. *F. prausnitzii* and its phylogroups abundance in inflammatory bowel disease patients depending on whether or not they have had intestinal resection during the course of the disease.

Diagnostics§	N	*F. prausnitzii**	p-value	Phylogroup I*	p-value	Phylogroup II*	p-value
UC							
non-resected	43	4.85±0.61	1.000	2.51±1.21	0.136	2.92±0.96	0.727
resected	1	4.91		3.45		2.68	
CD							
non-resected	41	4.86±1.43	0.016	1.52±1.84	0.379	2.11±1.46	0.001
resected	13	3.74±0.78		0.45±1.07		0.65±0.84	

* Median $\log_{10}$ 16S rRNA gene copies/ million bacterial 16S rRNA gene copies ± standard deviations
§UC, ulcerative colitis; CD, Crohn's disease

Table S7. *F. prausnitzii* and its phylogroups abundances (median $\log_{10}$ 16S rRNA gene copies/ million bacterial 16S rRNA gene copies ± standard deviations) in inflammatory bowel disease by medication at sampling.

Diagnostics§	N	*F. prausnitzii**	P-value	Phylogroup I*	P-value	Phylogroup II*	P-value
UC							
No treatment	25	4.95±0.65		2.51±1.32		2.93±1.03	
Mesalazine	6	5.02±0.33	0.904	2.53±0.84	0.806	3.31±0.98	0.832
Moderate immunosuppresants	9	4.56±0.58		2.75±0.41		2.85±0.71	
Anti-tumour necrosis factor	7	4.44±0.83		3.16±1.93		2.92±1.07	
CD							
No treatment	21	4.86±1.66		0.69±2.04		2.70±1.71	
Mesalazine	3	5.10±0.41	0.225	1.71±1.67	0.854	2.63±1.89	0.738
Moderate immunosuppresants	19	4.01±0.95		0.71±1.45		1.23±1.48	
Anti-tumour necrosis factor	16	4.01±1.43		0.67±1.48		1.49±1.18	

* Median $\log_{10}$ 16S rRNA gene copies/ million bacterial 16S rRNA gene copies ± standard deviations
§UC, ulcerative colitis; CD, Crohn's disease

Supplementary references

1. Lopez-Siles M, Khan TM, Duncan SH, et al. Cultured representatives of two major phylogroups of human colonic *Faecalibacterium prausnitzii* can utilize pectin, uronic acids, and host-derived substrates for growth. Appl Environ Microbiol. 2012;78:420-428
2. Barcenilla A, Pryde SE, Martin JC, et al. Phylogenetic relationships of butyrate-producing bacteria from the human gut. Appl Environ Microbiol. 2000;66:1654-1661
3. Cato EP, Salmon, Carolyn W, and W.E.C. Moore. *Fusobacterium prausnitzii* (Hauduroy et al.) Moore and Holdeman: emended description and designation of neotype strain. Int J Syst Bacteriol. 1974;24:225-229
4. Duncan SH, Hold GL, Harmsen HJ, et al. Growth requirements and fermentation products of *Fusobacterium prausnitzii*, and a proposal to reclassify it as *Faecalibacterium prausnitzii* gen. nov., comb. nov. Int J Syst Evol Microbiol 2002;52:2141-2146
5. Louis P, Duncan SH, McCrae SI, et al. Restricted distribution of the butyrate kinase pathway among butyrate-producing bacteria from the human colon. J Bacteriol. 2004;186:2099-2106
6. Lane DJ. 16S/23S rRNA sequencing. New York: In E. Stackebrandt and M. Goodfellow (ed.)., John Willy and Sons; 1991
7. Weisburg WG, Barns SM, Pelletier DA, et al. 16S ribosomal DNA amplification for phylogenetic study. J Bacteriol. 1991;173:697-703

In the last years, there has been a raising interest in *Faecalibacterium prausnitzii* and its role in maintaining intestinal health. In this PhD thesis *F. prausnitzii* populations of patients with gut disease and healthy individuals have been studied.

First, a phenotypic characterization of isolates from healthy volunteers has been performed, which has allowed to gain insight into the physiology of this species. A possible link between *F. prausnitzii* sensitivity to changes in gut physicochemical conditions and its disappearance in a diseased gut has been revealed. This relationship can serve as a basis for future studies aimed at recovering this microorganism population in the intestine.

Second, molecular studies on *F. prausnitzii* populations have been conducted. This has allowed to define two phylogroups within this species, and to describe the diversity of phylotypes in healthy individuals and in patients with intestinal disease. For the first time, which phylotypes are specifically compromised in patients suffering gut disorders such as ulcerative colitis, Crohn's disease and colorectal cancer have been identified.

Finally, these results have been used to design molecular tools which have been applied for the detection and quantification of this species and its phylogroups. These works are of great scientific importance and commercial interest because novel biomarkers are targeted, which in turn can be implemented as complementary molecular tools for the diagnosis and prognosis of intestinal diseases